



# **Análisis de la evaluación estatal de programas y proyectos en Colombia de 1994 a 2010**

**JOHANNA ALEJANDRA MORA RAMOS**

Universidad Nacional de Colombia  
Facultad de Derecho, Ciencias Políticas y Sociales, Departamento de Ciencia Política  
Bogotá, Colombia  
2015



# **Análisis de la evaluación estatal de programas y proyectos en Colombia de 1994 a 2010**

**JOHANNA ALEJANDRA MORA RAMOS**

Tesis de investigación presentada como requisito parcial para optar al título de:  
**Magíster en Políticas Públicas**

Director:

Ph.D. André-Noël Roth Deubel

Línea de Investigación:

Políticas Públicas

Universidad Nacional de Colombia

Facultad de Derecho, Ciencias Políticas y Sociales, Departamento de Ciencia Política

Bogotá, Colombia



## Resumen

En Colombia, la evaluación de políticas y programas es reciente y aunque ha evolucionado con la operación del Sistema Nacional de Evaluación de Resultados de la Gestión Pública -SINERGIA, los mecanismos y sistemas creados para ello se han orientado fundamentalmente a la medición de resultados de los principales programas del Plan Nacional de Desarrollo (PND). La presente investigación contiene una reconstrucción histórica y analiza la configuración y operación de la política pública de evaluación en Colombia, tomando como insumo clave la identificación de los referenciales de política y el estudio crítico de los informes finales de las evaluaciones de impacto producidos por el Departamento Nacional de Planeación, a través de un instrumento propio diseñado para ello, con base en estándares internacionales de calidad de la evaluación y retomando el abordaje metodológico de la metaevaluación, con el fin de reconstruir la historia de la política y comprender sus lógicas y formas de operación. Se concluye que la orientación a la medición de resultados es limitada y aún hay un margen importante para mejorar, especialmente de cara a la metaevaluación, la metodología, el contexto y el uso de las evaluaciones, y se recomienda elaborar una política de evaluación que responda a los niveles de complejidad de la sociedad colombiana y contribuya a la transparencia, la participación y el control ciudadano.

Palabras clave: Políticas públicas, evaluación, metaevaluación, medición de resultados.

In Colombia, policy evaluation is recent and even when the government has achieved a considerable success in the creation and strengthening of its whole-of-government monitoring and evaluation (M&E) system, formally the National System for Evaluation of Public Sector Performance (SINERGIA), the mechanisms created to evaluate, has been oriented to measure the inputs, outputs, outcomes and impacts of Presidential Goals, or vis-à-vis the national development plan. This dissertation build the policy referenciales and uses de metaevaluation approach to analyse the evaluation reports produced by Colombian Department of National Planning. The dissertation concludes among others, that public policy evaluation should be improved, in particular in terms of methodology and metaevaluation and recommends to ensure that evaluation information should be utilized more intensively.

Keywords: Public policy, evaluation, metaevaluation, measurement results.

## Contenido

<b>1. CAPÍTULO I: LA PRÁCTICA EVALUATIVA EN COLOMBIA .....</b>	<b>3</b>
1.1 BREVE RECONSTRUCCIÓN HISTÓRICA DE LA EVALUACIÓN EN COLOMBIA.....	14
1.2 EL SISTEMA NACIONAL DE MONITOREO Y EVALUACIÓN DE COLOMBIA: SINERGIA .....	15
1.3 POLÍTICA PÚBLICA DE EVALUACIÓN: UNA TRAYECTORIA DE REFERENCIALES ORIENTADOS A LA MEDICIÓN DE RESULTADOS .....	18
1.3.1 <i>Enfoque teórico y conceptual de los referenciales de política</i> .....	19
1.3.2 <i>Referenciales de la política: documentos estatales</i> .....	21
1.3.3 <i>Conclusiones</i> .....	27
<b>2. CAPÍTULO II: CRITERIOS PARA ANALIZAR LA EVALUACIÓN.....</b>	<b>29</b>
2.1 META-EVALUACIÓN .....	29
2.2 ESTÁNDARES DE EVALUACIÓN UTILIZADOS A NIVEL MUNDIAL .....	33
2.3 METODOLOGÍA: INSTRUMENTO DISEÑADO PARA EL ANÁLISIS DE LAS EVALUACIONES ESTATALES DE IMPACTO .....	43
2.4 SELECCIÓN DE SISTEMA DE CALIFICACIÓN Y DEFINICIÓN DE LA MUESTRA.....	47
2.4.1 <i>Sistema de calificación</i> .....	49
<b>3. CAPÍTULO III: ¿QUÉ NOS DICEN LAS EVALUACIONES? PRESENTACIÓN DE HALLAZGOS Y ANÁLISIS.....</b>	<b>53</b>
3.1 MARCO INSTITUCIONAL .....	54
3.2 ENFOQUE Y METODOLOGÍA.....	55
3.3 INFORME FINAL .....	56
3.4 IMPACTO DE LA EVALUACIÓN .....	58
<b>4. CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>59</b>
4.1 CONCLUSIONES .....	59
4.2 RECOMENDACIONES .....	59
<b>ANEXOS.....</b>	<b>61</b>
<b>BIBLIOGRAFÍA .....</b>	<b>108</b>

## Lista de acrónimos

<b>DNP</b>	Departamento Nacional de Planeación de Colombia
<b>DEPP</b>	Dirección de Evaluación de Políticas Públicas
<b>EeA</b>	Empleo en Acción
<b>FeA</b>	Familias en Acción
<b>FeA-PD</b>	Familias en Acción para población desplazada
<b>BID</b>	Banco Interamericano de Desarrollo
<b>Conpes</b>	Consejo Nacional de Política Económica y Social
<b>EI</b>	Evaluación de impacto
<b>HCB</b>	Programa Hogares Comunitarios de Bienestar
<b>IFS</b>	Institute for Fiscal Studies de Londres
<b>JeA</b>	Programa Jóvenes en Acción
<b>KEC</b>	Key Evaluation Checklist
<b>MEC</b>	Meta-Evaluation Checklist
<b>M&amp;E</b>	Monitoreo y Evaluación
<b>OCDE</b>	Organización de Cooperación y Desarrollo Económicos
<b>SIGOB</b>	Sistema de Gestión y Seguimiento a las metas del Gobierno
<b>SINERGIA</b>	Sistema Nacional de Evaluación de Resultados de la Gestión Pública
<b>VISU</b>	Programa de Vivienda de Interés Social Urbana





# Introducción

## 1.1. Antecedentes

En Colombia, la evaluación de programas y proyectos es reciente y aunque ha evolucionado en los últimos 16 años de operación del Sistema Nacional de Evaluación de Resultados de la Gestión Pública -SINERGIA, los mecanismos y sistemas creados para ello se han orientado fundamentalmente a la medición de resultados de los principales programas del Plan Nacional de Desarrollo (PND).

Esta orientación se ha inspirado en el paradigma científico positivista que privilegia los abordajes cuantitativos y emplea técnicas cada vez más sofisticadas de medición, que han dejado de lado otro tipo de aproximaciones cualitativas que permitirían cuestionar la política y no simplemente medir los resultados o impactos de los programas. Este enfoque, también encuentra su referente y su origen, en las directrices que sobre evaluación de políticas han establecido algunos organismos internacionales como la Banca Multilateral (Banco Mundial y Banco Interamericano de Desarrollo) y las asociaciones de académicos o expertos (Centro Latinoamericano de Administración para el Desarrollo -CLAD)

Adicionalmente, la evaluación realizada por el Estado colombiano ha estado a cargo del Departamento Nacional de Planeación (DNP) adscrito a la rama ejecutiva del poder público, por lo que se trata de una práctica de autoevaluación (el Estado evalúa lo que él mismo hace), que excluye la participación de la ciudadanía en su realización.

La investigación pretende reconstruir históricamente y analizar la configuración y operación de la política pública de evaluación en Colombia, para señalar los límites que implica su orientación, con el fin de realizar recomendaciones que permitan elaborar una política de evaluación que responda a los niveles de complejidad de la sociedad colombiana y contribuya a la transparencia, la participación y el control ciudadano.

La hipótesis planteada, sugiere que evaluación de políticas públicas que el estado colombiano ha desarrollado desde 1994 y hasta 2010, fue concebida como instrumento técnico de medición de resultados e impactos, a causa de la presión de organismos técnicos y de financiamiento internacionales y en consecuencia, se caracteriza por ser esencialmente cuantitativa y por privilegiar la medición de impactos y resultados de las políticas y programas más relevantes del PND. Además, debido a su alto nivel técnico es incierta su utilidad como herramienta para el control ciudadano y se asemeja a una práctica de autoevaluación estatal que parece estar más orientada a legitimar las decisiones del gobierno que a mejorar las políticas y programas nacionales y la gestión del Estado (3).

El análisis de la evaluación tomará como insumo clave la reconstrucción histórica y adicionalmente incluirá el estudio crítico de los documentos de evaluación producidos por el DNP, así como entrevistas semiestructuradas a algunos de los actores que han estado

y están involucrados en ella, con el fin de reconstruir la historia de la política y comprender sus lógicas y formas de operación.

Se espera que las conclusiones del estudio constituyan insumos valiosos para alimentar la política pública de evaluación y permitan ampliar su alcance y profundidad como herramienta necesaria para el control de la gestión del Estado en Colombia.

El objetivo general de esta investigación es conocer y analizar la configuración y operación de la política pública de evaluación en Colombia, desde 1994 y hasta 2010, para identificar sus fundamentos y definir sus alcances y limitaciones, con el fin de sugerir recomendaciones que permitan enriquecerla o transformarla.

De esta forma, se espera encontrar evidencia para refutar o confirmar la siguiente hipótesis: La evaluación de políticas públicas que el Estado colombiano ha desarrollado desde 1994 y hasta 2010, fue concebida como instrumento técnico de medición de resultados e impactos y en consecuencia, se caracteriza por ser esencialmente cuantitativa y por privilegiar la medición de impactos y resultados de las políticas y programas más relevantes para el gobierno del Plan Nacional de Desarrollo, por lo que se asemeja más a una práctica de autoevaluación estatal que parece estar orientada a legitimar las decisiones del gobierno que a mejorar las políticas y programas nacionales y la gestión del Estado.

La consecución de este objetivo general requiere responder las siguientes preguntas: i) ¿Cuáles fueron los referenciales de la política pública de evaluación que el Estado colombiano desarrolló entre 1994 y 2010? y ii) ¿Cuáles son las principales características de la evaluación que realiza el Estado colombiano a través del Departamento Nacional de Planeación?

# 1. Capítulo I: La práctica evaluativa en Colombia

El estudio de la política de evaluación estatal en Colombia es relativamente nuevo, debido a que la política misma es reciente en nuestro país y en la región. De las fuentes consultadas, se identifican distintos actores que estudian la evaluación de políticas, entre academia, organismos multilaterales, empresas consultoras privadas y el Estado.

Por lo general, se encuentran informes de diversas evaluaciones realizadas a políticas, programas y proyectos estatales, así como distintos manuales o textos de metodología para la evaluación de los mismos, pero son pocas las referencias que existen sobre el estudio de la evaluación de políticas, más aún para el caso colombiano. Quienes estudian la evaluación suelen ser académicos interesados en el tema, organismos multilaterales financiadores o co-financiadores de las evaluaciones, empresas privadas interesadas en la gestión pública o que realizan labores de consultoría, y algunas dependencias estatales encargadas o vinculadas a la evaluación. La limitada existencia de estudios sobre la evaluación de políticas en Colombia, constituye un aliciente y una justificación adicional para el presente estudio.

En este aparte se abordan 4 documentos que analizan la práctica evaluativa en Colombia desde distintas perspectivas.

El primer documento abordado, es un informe realizado en 2005 en virtud de un convenio de cooperación técnica entre el Banco Interamericano de Desarrollo y el Instituto de Investigación Económica Aplicada de Brasilia. El texto presenta una serie de estudios de caso sobre la función evaluación en algunos países de América Latina, incluida Colombia, basados en información oficial.

Su autor, Ricardo Toledo, sostiene que el origen de SINERGIA se sitúa en la Constitución Política de 1991, como ya se ha señalado, pero hace énfasis en dos aspectos fundamentales: primero, que la nueva Constitución surge en un contexto de gran conflicto y tensión social, por lo que estuvo orientada esencialmente a lograr la pacificación del país y el fortalecimiento del Estado colombiano, debilitado tras años de conflicto y a causa de la existencia de desigualdades sociales. Segundo, que desde su concepción, la evaluación en el país estuvo articulada al Plan Nacional de Desarrollo que se crea también en la nueva Constitución: “el SINERGIA fue concebido como contrapunto del Plan Nacional de Desarrollo que, como definido en su inserción legal, pasa a tener un papel central en el rescate de la gobernancia (sic) del país.

Es importante notar que este rescate se produce en un contexto singular de tensiones

políticas y conflictos armados que ponen en jaque las funciones más básicas del ejercicio del poder público y eso conduce al énfasis del Plan para un proceso de pacificación nacional largamente apoyado en políticas sociales con potencia de aglutinar segmentos sociales a su entorno. Sólo así es comprensible la estructura del Plan y, por consecuencia, la definición funcional del sistema de evaluación asociado.” (Toledo, 2005, p. 178).

De esta forma, Toledo llama la atención sobre el fuerte vínculo existente entre SINERGIA y los Planes Nacionales de Desarrollo (PND), lo cual debe tenerse presente en el estudio de la política de evaluación en Colombia como elemento constitutivo, pero además, el autor establece tres aspectos clave que orientan la planificación y, por esa vía, la evaluación en el país: “el SINERGIA es un sistema directamente relacionado a los planes nacionales de desarrollo establecidos por la Constitución de 1991. Donde se redefine el papel del Estado y amplía el significado constitucional de la planificación a partir de tres disposiciones: i) otorga al Plan Nacional de Desarrollo, carácter de ley con prioridad sobre los demás, volviendo indispensable la articulación entre los procesos de planificación y presupuesto; ii) establece los consejos de planificación como instancia consultiva en el proceso de aprobación del PND, obligando al Ejecutivo a reconocer al Congreso y a la sociedad civil como interlocutores indispensables en la definición de los objetivos prioritarios para el desarrollo del país; y iii) atribuye al Departamento Nacional de Planificación (DNP) la responsabilidad de proyectar y organizar los sistemas de evaluación que comparten los resultados de las políticas y proyectos de inversión” (Toledo, 2005, p. 176).

En segundo lugar, Toledo establece algunos límites del sistema de evaluación en Colombia, que se explican en parte debido a la estrecha relación entre Evaluación y PND: “del análisis de los subsistemas de evaluación previstos en el ámbito del SINERGIA, se observa la búsqueda de equilibrio entre la eficiencia y la eficacia en el conjunto de la acción pública. No obstante, por más que exista la intención explícita de que el proceso de evaluación sea dirigido primordialmente a la observancia de los principios y objetivos de desarrollo establecidos en la Constitución de 1991, esta conexión parece –en el diseño institucional descrito- mediada por una serie de condiciones que pueden volverla muy opaca. Siendo el PND considerado, desde el punto de vista del SINERGIA, el primer depositario de los principios y objetivos constitucionales de pacificación y reconstrucción social del país, el proceso de evaluación está contenido, por definición, bajo el techo del PND. Los grandes objetivos constitucionales son remitidos al PND y, en el ámbito del Plan, traducidos en programas y acciones. Lo que el SINERGIA puede evaluar, según la estructura descrita, es la conformidad de los programas y acciones respecto al Plan, e inclusive la conformidad de las entidades ejecutoras de estos programas y acciones, pero no su conformidad última a los objetivos nacionales.” (Toledo, 2005, pp. 191 - 192).

Este aspecto presentado por Toledo, se corresponde parcialmente con uno de los componentes de la hipótesis de trabajo, que señala el carácter limitado de medir impactos y resultados de políticas y programas del PND, pues en principio parece que la evaluación no tuviera mayores posibilidades de incidir en el diseño de las políticas. En adición, Toledo llama la atención sobre la imposibilidad de que la evaluación de políticas se extienda hasta la relación entre dichas políticas y lo que él denomina los “objetivos nacionales” haciendo referencia a lo establecido por la Carta Política de 1991.

En consecuencia, el autor concluye que “la estructura de evaluación montada en el SINERGIA tiende a mostrarse adecuada para fines de seguimiento y retroalimentación de las acciones del PND, pero presenta limitaciones si es confrontada con los objetivos más amplios de la acción del Estado y de su reforma a largo plazo. El principal límite está en la ambigüedad de la evaluación estratégica y en los hilos que se puedan establecer entre las metas de programas o sectoriales y los objetivos de la administración pública como un todo” (Toledo, 2005, p. 176).

Una vez abordado el SINERGIA como un todo, Toledo hace referencia a la evaluación en particular, y señala que, entre otros, “la evaluación y el desarrollo institucional de las políticas públicas del PND son orientados, en el ámbito del SINERGIA, por los siguientes criterios:

[...] que los procesos decisorios sean basados en las evaluaciones de resultados más que en el análisis de insumos; que se configure una administración pública de tipo contractual, en que burocracias más autónomas y flexibles respondan por la realización de resultados claramente definidos y [monitoreados]; que la sociedad civil cuente con mecanismos que permitan exigir una presentación de cuentas sobre los resultados de los procesos de toma de decisiones; [...]” (Toledo, 2005, p. 191). Es importante resaltar el énfasis que hace el autor en los resultados.

Toledo también llama la atención sobre el carácter de autoevaluación propio de SINERGIA y que se reproduce en los niveles locales de gobierno: “[...] la piedra angular del sistema continúa siendo el plan indicativo y su evaluación, esencialmente basada en un proceso de autoevaluación.” (Toledo, 2005, p. 192) “El proceso de evaluación local reproduce el esquema básico de los planes indicativos, en el sentido de ser altamente centrado en la autoevaluación.” (Toledo, 2005, p. 200).

Finalmente, Toledo señala un aspecto relevante y positivo que hasta el momento no había sido señalado y es que la evaluación de políticas en el país, no surge de manera desarticulada o independiente sino que hace parte de un esfuerzo más amplio de modernización estatal: “es cierto que la concepción del Proyecto SINERGIA y la propia estructura del PND son –en el plan conceptual y doctrinario- parte de un proceso más amplio de reforma del Estado que está en curso en el país” (Toledo, 2005, p. 201).

En conclusión, para el autor “independientemente de la evolución de los hechos, el SINERGIA –como sistema de evaluación oficial- tiene un alto grado de formalización, por ser acoplado al PND, que es uno de los componentes centrales de la acción del gobierno, con respaldo en la Constitución de 1991, con el objetivo mayor de la pacificación nacional. Este nivel de formalización no llega, sin embargo, a caracterizar la función evaluación como un proceso sistémico [...], puesto que el énfasis en la autoevaluación, que es funcional desde el punto de vista de la gestión de programas, no lo es desde la perspectiva de evaluación de las políticas públicas. Por otro lado, no son previstos procedimientos específicos de evaluación de políticas, inclusive en el ámbito de las llamadas evaluaciones estratégicas” (Toledo, 2005, pp. 201 - 202).

De esta forma, para Toledo también existe un carácter limitado en la política de evaluación en Colombia, determinado por su articulación al PND (que no obstante también contribuye a su formalización), la preeminencia de la autoevaluación y la falta de procedimientos de evaluación específicos de políticas públicas.

El segundo documento trabajado es escrito por André-Noël Roth y publicado en la Revista Reforma y Democracia del Centro Latinoamericano de Administración para el Desarrollo –CLAD, titulado “La evaluación de políticas públicas en Colombia: una mirada crítica a partir de las prácticas evaluativas oficiales de los programas de la “Red de Apoyo Social”” publicado en el año 2009.

En el texto se presenta un análisis de las evaluaciones publicadas por el DNP con ocasión de los programas de la Red de Apoyo Social, que surge en el año 2000 como parte de la Estrategia de Recuperación Económica y Social del Plan Colombia, diseñada para mitigar el impacto de la recesión económica sobre la población más vulnerable. La hipótesis del documento es que, a propósito de la evaluación, “el enfoque elegido no permite generar reales cambios sociales ni profundizar la democracia y la participación por su fuerte amarre a una concepción tecnocrática de la evaluación, que tiende a cumplir con una función de legitimación de las políticas del gobierno, alejándose así de su objetivo central” (Roth, 2009, p. 2).

Roth señala que existe una inclinación en las ciencias sociales a integrar el paradigma científico positivista como método de comprobación y fuente de legitimación, visible también en Colombia y en el ámbito de la evaluación oficial de políticas: “en particular, se trata de la fuerte tendencia al “managerialismo”, la difícil o imposible tenida en cuenta del pluralismo de los valores y la tendencia a privilegiar el paradigma de investigación científica tradicional” (Roth, 2009, p. 3).

Así pues, Roth sostiene que en el afán de dar un peso científico a su trabajo, los evaluadores en Colombia han recurrido a metodologías cuantitativas en la mayoría de los casos: “los evaluadores, con el deseo de quedar considerados como auténticos científicos, tienden a usar exclusivamente el paradigma positivista vigente en las ciencias naturales, lo que se refleja en particular en el uso de métodos cuantitativos sofisticados, con la pretensión de encontrar la “verdad” de forma inobjetable” (Roth, 2009, p. 4).

Por otro lado, el autor identifica a los evaluadores de los programas de la RAS, así como la metodología utilizada y concluye que “la mayoría de las evaluaciones contratadas por el DNP fueron con la misma empresa, conformada por el Institute of Fiscal Studies, IFS (un instituto de investigación inglés especializado en el análisis de las políticas sociales y económicas a partir de técnicas microeconómicas), Econometría (una empresa de consultoría privada colombiana especializada en estudios económicos) y otra empresa de consultoría nacional especializada en la recolección y tratamientos de datos. A nivel metodológico, los evaluadores han usado siempre la misma metodología cuasi experimental longitudinal (antes-después), que compara la situación de un grupo de beneficiarios con un grupo de control en distintos momentos. Las dos evaluaciones que fueron realizadas por otros evaluadores usaron una metodología similar. Como se nota, esta concepción de la evaluación se sitúa plenamente en la perspectiva científica tradicional” (Roth, 2009, p. 6).

Así pues, Roth critica el carácter unívoco de la política de evaluación y llama la atención en el hecho de que su utilidad es cuestionable, o en el mejor de los casos, imperceptible tanto para los ciudadanos como para los tomadores de decisiones: “claramente, las evaluaciones realizadas por el DNP, basadas en una ilusoria capacidad de medición objetiva con fines de alcanzar una óptima eficiencia técnico administrativa, no constituyen

herramientas e instrumentos para profundizar la democracia y favorecer una revisión de las políticas y programas desarrollados por el gobierno y las instituciones públicas. Tampoco, aparecen capaces de articular una diversidad de fuentes y voces presentes en la sociedad” (Roth, 2009, p. 13).

De esta forma, Roth insiste en cuestionar la utilidad de las evaluaciones y sugiere que éstas lo que terminan haciendo es legitimando las decisiones previamente tomadas: “la función central de las evaluaciones, y en particular las evaluaciones de tipo experimental, es aportar insumos para la toma de decisión. Sin embargo, en los dos casos estudiados se nota que la toma de decisión no se realizó con base en las evaluaciones presentadas. Al contrario, ¡parece ser que son los evaluadores quienes integran y legitiman en sus comentarios y recomendaciones las decisiones ya tomadas!” (Roth, 2009, p. 6) “Estos dos ejemplos parecen ilustrar a la perfección lo que se conocía desde ya un tiempo: el problema mayor de este tipo de evaluación es que los resultados se dan a conocer cuando ya no tienen mayor pertinencia política, y que, para tomar decisiones acerca de un programa, la importancia de los impactos observados no son necesariamente los elementos más centrales para fundamentar una decisión” (Roth, 2009, p. 12).

Concluye el autor que finalmente sus observaciones “confirman la sensatez de las críticas a las evaluaciones que se basan en un intento, siempre difícil y limitado en su validez, de evaluar las políticas y programas mediante el paradigma científico de la medición positivista, y en este caso con ninguna reflexión acerca de los valores, tal como lo plantea la evaluación de la tercera generación. Se construye una imagen positiva de los programas únicamente con base en la ausencia de política alternativa, en un contexto de extrema pobreza en donde cualquier esfuerzo generará probablemente algún efecto positivo. Por lo tanto, en este marco, las políticas y programas siempre resultan legitimados por la “ciencia” evaluativa” (Roth, 2009, pp. 12 - 13).

Hasta el momento, ambos documentos abordados conservan una perspectiva crítica frente a la evaluación oficial en nuestro país, pero el tercer estudio que se ha retomado es un diagnóstico realizado por la Oficina Independiente de Evaluación del Banco Mundial y que, contrario al documento de Toledo, destaca las labores del gobierno colombiano en materia de seguimiento y evaluación: “El gobierno de Colombia ha logrado un éxito considerable en la creación y fortalecimiento de su sistema de seguimiento y evaluación (SyE) a nivel de todo el gobierno, SINERGIA. Éste es uno de los más fuertes sistemas de SyE en América Latina, en lo que se refiere a tipos de SyE que realiza, su credibilidad promedio y su utilización (Mackay, 2007, p. iii).

En el diagnóstico de SINERGIA, Mackay insiste en dar un crédito importante a Álvaro Uribe, pues fue justamente durante sus dos periodos de gobierno que el seguimiento y la evaluación florecieron en Colombia: “la mayor parte de este progreso se ha dado desde el año 2002, después de la elección del Presidente reformista Álvaro Uribe.” (Mackay, 2007, p. iii). “Ciertamente, el rol del Presidente, como impulsor y usuario clave de tal sistema de seguimiento no tienen precedentes dentro de la región.” (Mackay, 2007, p. 23).

No obstante, uno de los aportes del estudio de Mackay es reconocer el peso de la banca multilateral en la configuración de la política de evaluación y de SINERGIA: “el origen del sistema fue decisión del Ministro de Hacienda con el fin de *reproducir en Colombia el enfoque del Banco Mundial respecto a la evaluación*. El gobierno identificó la estructura

básica para un sistema de SyE con la asistencia técnica del Banco Mundial. La primera fase de la evolución del sistema se realizó entre 1990 y 1996, la cual incluyó un requisito formal para evaluaciones en la Constitución revisada de 1991” (Mackay, 2007, p. 2, *énfasis agregado*).

Además de influenciar la orientación de la política de SyE, la banca multilateral es financiadora por excelencia de la misma: “hacia finales de este periodo, en el 2001, el Banco Mundial intensificó su apoyo a SINERGIA, no sólo a través de un nuevo Proyecto de Modernización de la Administración Financiera Pública (MAFP II) sino también copatrocinando, con el Banco Interamericano de Desarrollo, una serie de evaluaciones de impacto de dos grandes programas del gobierno: Empleo en Acción (un programa para la creación de nuevos puestos de trabajo) y Familias en Acción (un programa de transferencias monetarias condicionadas).” (Mackay, 2007, p. 2).

Así pues, “desde el año 2002, se han gastado 10.8 millones de dólares en SINERGIA, de los cuales casi la mitad corresponde al financiamiento otorgado por el BID, el 32% por el Banco Mundial, y el 8% por el USAID y el PNUD. Durante este periodo cuando había limitaciones fiscales macroeconómicas severas, el propio gobierno había financiado sólo el 12% de costos de SINERGIA. Evidentemente, este bajo nivel de apoyo financiero no es sostenible a largo plazo.” (Mackay, 2007, pp. 2 - 3).

A diferencia de lo planteado por Roth desde una perspectiva crítica, que pone en cuestión la utilidad de las evaluaciones en el proceso de toma de decisiones, para Mackay “el sistema de SyE también incluye una ambiciosa agenda de evaluaciones, particularmente en evaluaciones de impacto, y en los próximos años completarán un número cada vez mayor. Estas evaluaciones ya están comenzando a ser utilizadas como un elemento en la toma de decisiones del gobierno.” (Mackay, 2007, p. iii). Sin embargo, no presenta evidencias que sustenten esta afirmación.

Ahora bien, sobre las evaluaciones, Mackay sostiene que “constituyen el segundo componente principal de SINERGIA. Actualmente se han completado o se están llevando a cabo cerca de 16 evaluaciones y hay otras 17 que van a ser realizadas en el transcurso de los próximos cinco años. La mayoría de estas evaluaciones son evaluaciones de impacto y *comprenden un sofisticado análisis estadístico* de los beneficiarios del programa y de los grupos de control o comparación. También incluyen generalmente un foco en eficacia operacional y otros temas de gerencia. Una pequeña cantidad de otros tipos de evaluación están realizados, y éstos tienden a ser enfocados estrictamente en temas de gerencia e institucionales.” (Mackay, 2007, p. 9, *énfasis agregado*).

Por su parte, el mismo autor reconoce los altos costos que implican las evaluaciones, haciendo énfasis nuevamente, en que el principal aportante es la banca multilateral: “desde el año 2002, por lo menos dos tercios del total gastado en SINERGIA, que proviene de todas las fuentes –principalmente del BID, del Banco Mundial y del Gobierno – han sido gastados en evaluaciones. Su costo varía desde US\$15,000 hasta US\$2 millones para las evaluaciones de impacto rigurosas del componente rural de Familias en Acción” (Mackay, 2007, p. 9).

De esta forma, el diagnóstico de Mackay adscrito al Banco Mundial y su Oficina Independiente de Evaluación, se muestra complacido con el desarrollo de SINERGIA, pues es coherente con las directrices del Banco sobre evaluación y en esa medida, no



presenta ninguna crítica al sistema en sí mismo, sino que plantea algunos desafíos que deben ser superados como avanzar hacia una mayor institucionalización o hacia el logro de mayor autonomía de la política en términos financieros.

Un segundo estudio que resulta favorable a la política de evaluación en Colombia, es realizado por Julio Villareal en el marco de un convenio de cooperación entre el Banco Mundial y el CLAD.

Villareal llama la atención sobre tres elementos propios de la política de evaluación en Colombia: el énfasis en rendición de cuentas, la planeación estratégica y la búsqueda de conmensurabilidad de las acciones del gobierno (medición de resultados): “en esa medida, dentro de la evolución que ha tenido el sistema a lo largo de los años, se le ha dado un peso importante siempre al tema de la rendición de cuentas como un elemento que permite activar de manera efectiva el control político y la participación ciudadana. Adicional a ese elemento de rendición de cuentas también hay un interés fuerte por informar las decisiones políticas. Esto supone trabajar con base en una planeación estratégica que incluya metas y objetivos establecidos. En definitiva, lo que se pretende es buscar que esos planes y programas de gobierno se conviertan en acciones y resultados concretos que puedan ser medibles en el transcurso de su ejecución.” (Villareal, 2007, p. 49).

Sin embargo, a diferencia de Mackay, Villareal presenta algunas de las falencias que aún tienen el sistema y la política: “para ejemplificar, está el caso de la rendición de cuentas, en el cual SINERGIA efectivamente ha hecho un énfasis importante. Los resultados de esta orientación han sido la utilización de SINERGIA como una herramienta del Ejecutivo para hacer seguimiento y control de las acciones del mismo Ejecutivo. Esto no se ha traducido en una rendición de cuentas hacia fuera más contundente. No se puede negar que hay avances al respecto, como los consejos comunales o los consejos televisados de ministros, pero aún no es claro cómo la ciudadanía puede utilizar y usa efectivamente la información de SINERGIA para hacer valer la rendición de cuentas de sus gobernantes.” (Villareal, 2007, p. 50).

De otra parte, al igual que Mackay, Villareal enfatiza la importancia del gobierno pasado en la consolidación y fortalecimiento de la política de evaluación y de SINERGIA en particular: “no se puede negar que el impulso político que le ha dado la actual administración al sistema ha sido relevante para su consolidación. Ha sido una coyuntura favorable de la cual SINERGIA ha sabido beneficiarse; pero ahora el reto es darle permanencia a ese interés y asegurar que, sin importar la administración de turno, el conjunto de la administración pública vea la relevancia de utilizar la información de monitoreo y evaluación en la toma de decisiones y en la rendición de cuentas sobre el uso adecuado de los recursos y la efectividad de la gestión pública.” (Villareal, 2007, p. 51).

Una mirada a los documentos producidos con ocasión del análisis de la política de evaluación y particularmente de SINERGIA, permite concluir que la evaluación de políticas en Colombia ha avanzado notablemente desde su concepción, particularmente durante la administración de Álvaro Uribe. También, que el fundamento metodológico y epistemológico de la evaluación en el país es el paradigma científico positivista (tradicional), lo que se traduce en una metodología de evaluación predominante, cuantitativa y orientada a la medición de resultados. Asimismo, la utilidad de las

evaluaciones en los procesos de toma de decisiones y para el control social, aparece aún muy débil y puesta en cuestión. Finalmente, la influencia de la banca multilateral en la política de evaluación queda expuesta de manera clara, no sólo en términos de financiación sino también – y más importante aún- en términos de la orientación de la práctica evaluativa en Colombia.

Ahora bien, desde una perspectiva internacional, el campo de la evaluación apareció en los años sesenta, “cuando fue evidente que la investigación de la evaluación estaba dominada por una sola metodología, y los investigadores de la evaluación creían que su potencial era limitado” (Palumbo & Nachmias, 1998, p. 286) House (1998) señala que “la metodología que caracterizó la evaluación hasta inicios de los setenta estaba orientada a determinar si las metas se estaban alcanzando y se creía que [las evaluaciones] serían utilizadas de forma automática para mejorar los procesos sociopolíticos” (House, E., 1998, p. 287)

Para Susana Hintze (2001), “en términos de su orientación general, los paradigmas tradicionales [de evaluación] aparecen ligados al positivismo, a la racionalidad organizativa y técnica y a la neutralidad valorativa” (p. 5). Esta visión limitada de la evaluación ha permanecido en el tiempo, aún cuando han surgido otros enfoques y definiciones (paradigmas “alternativos” o “pluralistas”): “conversely, evaluation is traditionally seen as a postimplementation activity, a procedure for determining the effectiveness of a program after it has been operationalized.” (Palumbo, 1982, p. XI).

Sin embargo, en sintonía con lo señalado por Weiss, “lo que tienen en común todos los empleos del término [evaluación] es la noción de juzgar el valor o los méritos de alguna cosa. Alguien está examinando y sopesando un fenómeno (una persona, una cosa, una idea) valiéndose de algún patrón, explícito o implícito” (Weiss, 1999, p. 13).

Lincoln Guba e Yvonna Lincoln (1989) realizan un recuento sobre los cambios del significado de la evaluación durante cien años de historia y los clasifican en cuatro generaciones.

La primera es la de medición (measurement generation), en la cual la evaluación reviste un carácter eminentemente técnico (technical sense of evaluation) y el evaluador espera conocer todos los instrumentos disponibles de forma que cualquier variable pueda ser medida: “The role of evaluator was technical” (Guba & Lincoln, 1989, p. 26).

La segunda generación (description) está caracterizada por describir los patrones de fortaleza o debilidad respecto a los objetivos establecidos. El rol del evaluador es describir, aunque conserva su carácter técnico (Guba & Lincoln, 1989, p. 28). Sin embargo, la evaluación así definida carece de un elemento clave que la tercera generación (judgement), tratará de integrar: el juicio. Así pues, esta generación incluye el juicio como elemento propio del acto de evaluar y el rol de evaluador consiste en emitir juicios, conservando su labor descriptiva y técnica, heredada de las generaciones anteriores.

Finalmente, ante los límites de la evaluación de cada una de las generaciones, los autores proponen un enfoque alternativo: responsive constructivist evaluation, en el que los parámetros y límites de la evaluación no son decididos entre el evaluador y el cliente como en las tres primeras generaciones, sino que se determinan en un proceso de

negociación entre las partes involucradas; esta etapa consume la mayor cantidad de tiempo y recursos. (Guba, Lincoln, 1989, p. 39) La metodología constructivista, como su nombre lo indica, se sustenta en un enfoque alternativo al paradigma científico positivista que había sido un elemento continuo en las tres generaciones anteriores. Además, la cuarta generación de evaluación no indaga por los objetivos, las variables o las decisiones, sino que le interesa conocer los acuerdos (claims), las críticas (concerns) y los desacuerdos (issues). De ahí, que contenga cuatro etapas coherentes con un método dialéctico y hermenéutico.

La periodización propuesta por Guba y Lincoln, es coherente con el desarrollo epistemológico y empírico de la evaluación y Susana Hintze, sintetiza parte del debate entre enfoques tradicionales y alternativos de evaluación en 5 aspectos esenciales:

a) En términos de su orientación general, los paradigmas tradicionales aparecen ligados al positivismo, a la racionalidad organizativa y técnica y a la neutralidad valorativa. Los paradigmas alternativos se nutren de vertientes fenomenológicas e incorporan al análisis, el papel de los valores y la perspectiva del actor.

b) En lo que hace al rol del evaluador, en el primer caso éste es visto como el depositario de saberes técnicos y responsable del rigor científico. Para los paradigmas alternativos se acerca más a un mediador o negociador con capacidad de gestión política, que promueve la discusión inmerso en el propio programa.

c) En cuanto a las características metodológicas, los paradigmas tradicionales importaron del análisis económico los modelos de costo-beneficio, los que transitron después a los de costo-efectividad. Frente a ellos los alternativos se interesan más por el estudio de las necesidades sociales y la “perspectiva de los beneficiarios”. En lo que respecta al tipo de diseño, las técnicas y la información utilizada, los paradigmas tradicionales desarrollan modelos bastante complejos de tipo experimental o cuasi experimental apoyados en técnicas estandarizadas de producción de información cuantitativa de base probabilística que requiere para su análisis de procedimientos estadísticos. Los paradigmas alternativos en oposición -en algunos casos en explícito rechazo a éstos- se inclinan por diseños eclécticos específicamente definidos en función de las características de los programas a evaluar, con información cuali-cuantitativa generada por encuestas, entrevistas, observación y estudios de caso.

d) En lo que refiere a los resultados de la evaluación, los paradigmas tradicionales parten de objetivos y metas y analizan procesos e impactos. Los estudios enmarcados en este enfoque presentan una fuerte preocupación por la calidad técnica de las evaluaciones, la validez de los resultados y la rigurosidad metodológica. Los alternativos se preocupan más por relacionar efectos con procesos y resultan mucho más pragmáticos en términos de proveer respuestas concretas a necesidades inmediatas de gestión. (Hintze, 2001, p. 5)

Para terminar, Susana Hintze, con ocasión de la evaluación de programas sociales, refiere una serie de características del “deber ser” de la evaluación, con las que parecen estar de acuerdo académicos, evaluadores y técnicos y que se retoma en esta investigación:

- La evaluación es una herramienta para reformar políticas y un mecanismo de mejoramiento y transformación de los programas sociales, en ese sentido debe detectar

tanto lo que funciona imperfectamente como destacar lo que marcha bien.

- No puede ser considerada como un fin en sí mismo, sino un instrumento para el desarrollo de programas y proyectos sociales: configura un insumo para la toma de decisiones sobre políticas y asignación de recursos.
- Para lograrlo debe ser desmitificada de los contenidos que le han sido atribuidos y del uso que se le ha dado en muchos casos (panacea para todos los problemas; carácter amenazante; instrumento de control político y de distribución de premios y castigos; mecanismo de racionalización del gasto, etc.).
- Constituye un tipo particular de investigación social que demanda de amplitud y flexibilidad metodológica y técnica y debe generar información confiable y adecuada a la gestión.
- La evaluación es un proceso de aprendizaje a través del cual es posible resolver en conjunto los problemas que afrontan los programas sociales.
- Requiere ser institucionalizada a través de una cultura de la evaluación y constituirse en una herramienta de la gestión integral de las políticas sociales. (Hintze, 2001, p. 4)

Aunque se conservan en la actualidad muchos de los elementos de la visión clásica de evaluación, se evidencia un llamado de algunas corrientes a que la evaluación de las políticas públicas trascienda su sentido instrumental como herramienta de medición, y avance hacia la construcción de nuevas ideas, metodologías y herramientas.

Desde su concepción en la Constitución de 1991, la evaluación de políticas públicas en el país estuvo orientada a medir resultados e impactos y en ese sentido se insistió en la idea de que constituía un instrumento técnico, lo cual determinó las herramientas, usos y tipos de evaluación que ha empleado el Estado hasta la fecha. Además, se atribuyó al Estado mismo, en el organismo de planeación, la responsabilidad de evaluar:

“Artículo 343. La entidad nacional de planeación que señale la ley, tendrá a su cargo el diseño y la organización de los sistemas de evaluación de gestión y resultados de la administración pública, tanto en lo relacionado con políticas como con proyectos de inversión, en las condiciones que ella determine. Artículo 344. Los organismos departamentales de planeación harán la evaluación de gestión y resultados sobre los planes y programas de desarrollo e inversión de los departamentos y municipios, y participarán en la preparación de los presupuestos de estos últimos en los términos que señale la ley. En todo caso el organismo nacional de planeación, de manera selectiva, podrá ejercer dicha evaluación sobre cualquier entidad territorial.” (Constitución Política de Colombia; 1991: Artículos 343 y 344)

Dicha conceptualización de la evaluación, que parece aproximarse a los paradigmas tradicionales a los que se refiere Hintze, no surgió de manera espontánea sino que obedeció a dinámicas propias de la lucha política que finalmente la establecieron como dominante y, conocer, caracterizar y analizar la forma como se construyó, se adoptó o se copió, permite entender por qué la política se ha desarrollado de una forma particular.

Es así, que desde su concepción en la Constitución Política de Colombia de 1991, que

sólo menciona la evaluación en dos momentos –en el capítulo II referido a los planes de desarrollo, en los artículos 343 y 344-, la evaluación se articula a la gestión y los resultados de la administración pública, y contempla dos dimensiones, la de políticas y la de proyectos de inversión.

Posteriormente, en el decreto 2167 de Diciembre 30 de 1992 , en el capítulo I referido a estructura y funciones, artículo 2, se determinan entre las funciones del Departamento Nacional de Planeación “elaborar las metodologías para la preparación y evaluación de los programas y subprogramas del Plan Nacional de Desarrollo, así como para la identificación, formulación y evaluación de los proyectos financiados con los recursos del Presupuesto General de la Nación; 3. Diseñar y organizar sistemas de evaluación de gestión y de resultados de la administración, tanto en lo relacionado con políticas como con proyectos de inversión, y señalar los responsables, términos y condiciones para realizar la evaluación.

En este artículo, vigente hasta el año 2000, nuevamente se define la evaluación articulada a programas, subprogramas y proyectos, pero ahora acota aún más su espacio al circunscribirlos al Plan Nacional de Desarrollo y a la característica de ser financiados con recursos del Presupuesto General de la Nación.

El artículo 2 del mismo decreto establece las atribuciones de las dependencias en relación con el Plan Nacional de Desarrollo y señala que las atribuciones de las dependencias que conforman la estructura del Departamento Nacional de Planeación, se ejercerán en función de las actividades de elaboración, formulación, evaluación y seguimiento del Plan Nacional de Desarrollo, así como de las políticas, estrategias, planes, programas y proyectos aprobados por el Conpes, sin perjuicio de otras disposiciones legales sobre atribución de competencias.

Aunque aparentemente este artículo aumenta el número de elementos objeto de la evaluación, continúa acotándolos cada vez más, pues construye diversas categorías que terminan articulando la evaluación al PND, las políticas, estrategias, planes, programas y proyectos aprobados por el Conpes.

Finalmente, el artículo 22 de este decreto que habla sobre la División especial de evaluación y control de gestión, en los numerales 2 y 5, señala como sus funciones “Diseñar, homologar y difundir entre las entidades públicas, las metodologías de control de gestión y evaluación de resultados de las políticas, planes, programas y proyectos ejecutados con recursos públicos”, así como realizar la evaluación de gestión y resultados sobre los planes y programas de las entidades territoriales, que indique el Director del Departamento. Aquí se evidencia la tipificación de la evaluación como evaluación de resultados de políticas, planes, programas y proyectos, ejecutados con recursos públicos. Esta definición implica de hecho la exclusión de otros tipos u objetivos de la evaluación y la circunscribe exclusivamente a los resultados.

Al respecto, André-Noël Roth (2009), en su análisis de las evaluaciones oficiales realizadas a los distintos programas de la Red de Apoyo Social (programa que surge en el año 2000 para hacer frente a los efectos sociales de la crisis económica que afrontaba el país), sugiere que “la práctica evaluativa colombiana “oficial” presente sigue centrándose, a la vez que contribuye a su reforzamiento, en una alianza del saber científico exclusivo en manos de expertos con un poder político fundamentalmente

concebido desde una perspectiva autoritaria y vertical.” (Roth, 2009, p. 13)

A partir del estudio de dichas evaluaciones concluye que se encuentran en clara sintonía con el paradigma positivista (tradicional) del que habla Hintze: “estas observaciones confirman la sensatez de las críticas a las evaluaciones que se basan en un intento, siempre difícil y limitado en su validez, de evaluar las políticas y programas mediante el paradigma científico de la medición positivista, y en este caso con ninguna reflexión acerca de los valores, tal como lo plantea la evaluación de la tercera generación. (Roth, 2009, p. 12)

## **1.1 Breve reconstrucción histórica de la evaluación en Colombia**

La práctica evaluativa en el país tuvo dos momentos clave entre 1994 y 2010. El primero de ellos ocurre en 1994 con la creación del Sistema Nacional de Evaluación de Resultados de la Gestión Pública (Sinergia) y el segundo se da en 2004 durante la administración del presidente Álvaro Uribe.

En la Constitución Política de 1991 se estableció la necesidad de crear un sistema de evaluación para el nivel nacional que ayudara a medir los resultados de las políticas y proyectos implementados por el Gobierno Nacional. Se ordenaba concretamente la organización de un sistema de evaluación de gestión y resultados en el sector público en cabeza de la entidad nacional de planeación. En cumplimiento de lo anterior, el Decreto 2167 del 30 de diciembre de 1992, estableció que al Departamento Nacional de Planeación le correspondía “diseñar y organizar sistemas de evaluación de gestión y de resultados de la administración, tanto en lo relacionado con políticas como proyectos de inversión, y señalar los responsables, términos y condiciones para realizar la evaluación.”

En consecuencia, en el año 1994 se creó Sinergia, con el propósito de mejorar la efectividad en la formulación y ejecución de las políticas del Plan Nacional Desarrollo (PND) y medir en el momento y de manera apropiada los resultados de la gestión pública, y se constituyó en herramienta indispensable para la rendición de cuentas del Gobierno nacional ante la ciudadanía<sup>1</sup>.

Durante los primeros años de operación de Sinergia, el énfasis estuvo en la institucionalización del Sistema y en el diseño robusto del mismo. Sin embargo, la práctica evaluativa no constituyó una prioridad hasta 2002 cuando el presidente Álvaro Uribe, da un impulso determinante a la práctica de la evaluación, particularmente de la evaluación de impacto, pues fue justamente durante sus dos periodos de gobierno que el

---

<sup>1</sup> El Sistema Nacional de Evaluación de Gestión y Resultados (SINERGIA), nace en 1994 con el ánimo de “introducir una cultura de la evaluación al interior de las entidades con el objeto de fortalecer la capacidad de manejo gerencial de la inversión pública” (CONPES 2688, 1994) Posteriormente, sus objetivos se redefinen en documento CONPES 3294 de 2004.

seguimiento y la evaluación florecieron en Colombia. Mackay (2007: iii) reconoce la importancia del papel desempeñado por el expresidente (impulsor y usuario clave del sistema de seguimiento y evaluación) como un aspecto esencial y una característica propia del caso colombiano.

En efecto, el expresidente institucionalizó una práctica de rendición de cuentas a la ciudadanía a través de los Concejos Comunales de Gobierno realizados semanalmente en distintos lugares del país. En estos espacios el exmandatario presentaba los avances y resultados de la gestión de su gobierno y fue justamente esta práctica la que terminó dando un impulso a Sinergia, pues de este sistema se obtenía la información necesaria.

Posteriormente, en 2004 con el Documento CONPES 3294, se formalizaron los avances logrados durante la etapa de diseño e implementación de Sinergia (1994-2004), y se definió su campo de acción a partir de tres componentes: i) seguimiento al PND, ii) evaluaciones de políticas estratégicas y iii) rendición de cuentas, los cuales desembocaron en el desarrollo del Sistema de Seguimiento a Metas del Gobierno (Sismeg, ahora Sinergia Seguimiento) y el Sistema Nacional de Evaluaciones (Sisdeval, ahora Sinergia Evaluaciones) así como en la elaboración, de manera concertada, de un Manual de Rendición de Cuentas y un índice que califica a las entidades del Gobierno de acuerdo a la calidad de sus rendiciones (DNP, 2012: 9).

En este contexto, y particularmente en el segundo periodo presidencial, se empezaron a realizar evaluaciones de impacto en el país que, dicho sea de paso, fueron sobre algunos de los programas más importantes del gobierno de turno.

Así pues, en Colombia, la evaluación de programas y proyectos es reciente y aunque ha evolucionado en los últimos 21 años de operación de Sinergia, los mecanismos y sistemas creados para ello se han orientado fundamentalmente a la medición de resultados de los principales programas del PND.

## **1.2 El Sistema Nacional de Monitoreo y Evaluación de Colombia: SINERGIA**

El Sistema Nacional de Evaluación de Resultados de la Gestión Pública busca mejorar la efectividad en la formulación y ejecución de las políticas, Programas y proyectos públicos, así como, medir oportunamente los resultados de la gestión de los administradores públicos y, fortalecer el manejo gerencial de la inversión pública y generar la información adecuada para la asignación de recursos (Conpes 3294/2004).

El Sistema tiene dos componentes: el Sistema Nacional de Evaluaciones –Sisdeval- y el Sistema de Seguimiento a Metas de Gobierno –Sismeg-

Sisdeval que es el que nos ocupa en el presente estudio, está conformado por cuatro instancias:

El Consejo Nacional de Política Económica y Social –Conpes- como la instancia máxima de orientación y decisión del Sistema.

El Comité Intersectorial de Evaluación y Gestión por Resultados como ámbito de coordinación técnica y de vinculación entre el presupuesto y la evaluación.

El Departamento Nacional de Planeación como coordinador técnico del Sistema, a través de la Dirección de Evaluación de Políticas Públicas –DEPP-.

Las entidades cabeza de sector, del sector descentralizado y entidades territoriales objeto de seguimiento y evaluación.

**Diagrama 1-1** Componentes del Sistema Nacional de Evaluaciones –Sisdeval



Fuente: Elaboración propia.

El Comité Intersectorial de Evaluación de Gestión y Resultados es creado mediante la Ley 1151 de 2007, en el artículo 132, de conformidad con lo dispuesto en el artículo 29 de la Ley 152 de 1994, en relación con el diseño y organización de los sistemas de evaluación de gestión y de resultados, como una instancia para:

Establecer las prioridades de evaluación de programas.

Aprobar las metodologías de evaluación y su alcance.

Considerar los resultados de las evaluaciones que puedan contribuir a mejorar los procesos de formulación de políticas.

Hasta 2008, el rol del comité se había centrado en la evaluación de temas sociales, la agenda de evaluación se definía por la demanda de entidades que tenían recursos y la agenda se compuso principalmente de evaluaciones de impacto.



En el 2009, la Dirección de Evaluación Políticas Públicas del DNP propuso la ampliación del CIE para incrementar la cobertura de temas evaluables, implementar un esquema balanceado de oferta y demanda de evaluaciones, fortalecer la implementación de las recomendaciones de las evaluaciones, incorporar otras metodologías (e.g. Evaluaciones Ejecutivas – E2) e institucionalizar Comités Sectoriales

**Cuadro 1-1** Roles del Comité Intersectorial de Evaluación

Hasta 2008	2009 y en adelante
<p>El rol del comité se ha centrado en promover la evaluación de temas sociales.</p> <p>La agenda se ha compuesto principalmente de evaluaciones de impacto.</p> <p>La agenda de evaluación se define por la demanda de entidades que cuentan con recursos.</p>	<p>Incrementar la cobertura de temas de evaluación en términos de los objetivos de política sectorial.</p> <p>Implementar un esquema balanceado de oferta y demanda de evaluaciones.</p> <p>Garantizar que las recomendaciones de las evaluaciones se conviertan en planes de mejoramiento (e.g. hogares comunitarios de bienestar).</p> <p>Incentivar el uso de las metodologías de evaluación de acuerdo a los requerimientos de información del sector.</p> <p>Institucionalizar los Subcomités Sectoriales de M&amp;E para la formulación de planes de trabajo estratégicos enfocados a una gestión más eficiente del sector.</p>

Fuente: DNP, 2011

Ahora bien, la construcción de las agendas de evaluación se compone de cuatro fases:

1. Elaboración de un compendio de las principales temáticas y estrategias desarrolladas por las entidades para implementar las políticas sectoriales.
2. Priorización de temas evaluables asociados a los principales requerimientos de información para cada una de los programas.
3. Identificación de los requerimientos de información (preguntas a resolver) y alcances del proyecto de evaluación.
4. El Ministerio como cabeza de sector consolida la propuesta del sector.
5. La identificación de las herramientas de evaluación más pertinentes para dar respuesta a los requerimientos de información.
6. Presentación de las propuestas de evaluación sectoriales al CIE Ampliado.

Adicionalmente, el sistema contempla una estrategia de difusión que consiste en un proceso que tiene tres instancias de discusión que deben surtirse en estricto orden.

- Primera instancia: Una vez finalizada la evaluación, se hace una sesión de discusión con el equipo de la consultoría y con las contrapartes técnicas de gobierno. Tiene como objetivo revisar los resultados de la evaluación, discutir su viabilidad operativa y aclarar las dudas existentes entorno a las recomendaciones sugeridas por el proyecto de evaluación.
- Segunda instancia: Presentar a las cabezas del sector y/o director del programa los resultados de la evaluación. Esta sesión debe estar acompañada por el director técnico o subdirector técnico de la Dirección que corresponda en DNP y demás entidades.
- Tercera instancia: Presentar en los Comités Sectoriales las recomendaciones y hallazgos de la evaluación. Con el fin de garantizar que la evaluación derive en un Plan de Mejoramiento del Programa evaluado.

Como se observa, la función de evaluación se encuentra en cabeza del Departamento Nacional de Planeación y es uno de los componentes de Sinergia y entre 1994 y 2010, tenía actores, roles y procesos establecidos.

Ahora bien, en la práctica de la evaluación estatal en Colombia no se evidencia su utilidad para el control ciudadano, pues el componente de rendición de cuentas está articulado al de Seguimiento de metas de gobierno –SIGOB- y aparece separado de Sistema Nacional de Evaluaciones –SISDEVAL-, sin que el vínculo entre ambos resulte claro, es de rescatar que la DEPP en el año 2010 participó activamente en la formulación de la Política de rendición de cuentas de la rama ejecutiva a los Ciudadanos, cuyo marco está en el documento CONPES 3654 de 2010 y que tiene por objeto “ordenar y articular las diferentes acciones que en el país se han venido presentando en materia de rendición de cuentas con un marco conceptual y una estrategia amplia que permita garantizar la sinergia de estas acciones en pro de ofrecer a la ciudadanía unas mejores y más claras explicaciones sobre el actuar del sector público” (Pág. 4)

Es importante señalar que en 2012 Sinergia fue reestructurado, por lo que la estructura presentada en este capítulo corresponde a la que estaba vigente durante el periodo que abarca el presente estudio (1994 - 2010).

Este aparte tenía como objetivo esbozar de manera general el arreglo institucional en el que se encuentra inmersa la función de evaluación en el país, a manera de contextualización y por lo tanto no pretende ser exhaustiva.

En el siguiente numeral se identifican y analizan de forma detallada los referenciales que operaron en la formulación de la política de evaluación de Colombia, con el fin de responder a la primera pregunta de investigación.

### **1.3 Política Pública de Evaluación: una trayectoria de referenciales orientados a la medición de resultados**

A primera vista, el Estado colombiano ha tenido importantes avances en materia de evaluación de políticas públicas desde los años noventa, lo cual se ha materializado en la creación de herramientas y sistemas útiles para evaluar la gestión y las políticas de

gobierno. Sin embargo, es necesario analizar los referenciales que determinaron las directrices de la política de evaluación desde su origen en 1991 y hasta 2010, pues esto permitirá conocer y analizar la manera como se configuraron y se llevaron a cabo las diversas acciones emprendidas por el Estado en la materia.

Este capítulo trata precisamente de contribuir en este análisis y tiene como objetivo describir y analizar los referenciales de la política pública de evaluación en Colombia, que se inscribe al fenómeno de la “evaluación de las políticas públicas” y sus resultados, a la luz de la perspectiva de abordaje de las políticas públicas como campo cognitivo.

La hipótesis de trabajo formulada para este análisis indica que los referenciales de la política pública de evaluación en Colombia desde 1991 hasta la actualidad, han sido determinantes para que la orientación de dicha política corresponda principalmente a la medición de resultados.

Los componentes privilegiados para ser estudiados son los referenciales de política utilizados en el proceso de diseño e implementación de la política pública de evaluación de políticas, a partir de su uso en categorías que definen directrices, contenidas en documentos estatales.

En este marco, se indaga por qué y cómo los referenciales de la política pública de evaluación en Colombia entre 1991 y 2010, fueron determinantes para que la orientación de dicha política se inscriba principalmente a la medición de resultados.

A efectos de la exposición, el capítulo se organiza en tres partes. Primero, se encuentra un aparte introductorio de naturaleza teórica que presenta la conceptualización básica, que constituye el referente de trabajo articulado a los referenciales de política y al abordaje de las políticas públicas como campo cognitivo.

El segundo aparte, atendiendo al problema de investigación y al enfoque teórico, se ocupa de identificar y analizar los referenciales de política, a partir del uso de conceptos y paradigmas, en los documentos que dan cuenta del diseño y la implementación de la política pública de evaluación.

Finalmente, se presentan las conclusiones que recogen los hallazgos y verificaciones realizadas a lo largo del capítulo.

### **1.3.1 Enfoque teórico y conceptual de los referenciales de política**

El propósito de este aparte es presentar el marco conceptual básico sobre los referenciales de política, que se circunscribe al fenómeno de las políticas públicas y la configuración de formas de intervención de Estado. Como se ha dicho, se analizan particularmente los referenciales operados en el diseño y ejecución de distintas acciones de la evaluación de políticas públicas en el país, que constituyen procesos de naturaleza estatal, que operacionalizan parte de las acciones de política del Departamento Nacional de Planeación sobre evaluación de la gestión pública y de sus resultados.

En ese sentido, cabe señalar que a efectos del presente análisis, la política pública será entendida como campo cognitivo y de representaciones, a partir del abordaje del concepto de referenciales de política propuesto por Pierre Muller (2008), sugiriendo que dichos referenciales operan estructuras categoriales que determinan formas de intervención estatal concretas en el ámbito de la evaluación, orientadas principalmente a la medición de resultados.

Desde una perspectiva tradicional de análisis de las políticas públicas, se ha partido de la idea según la cual, su origen y concepción revisten un carácter fundamentalmente técnico. Sin embargo, Pierre Muller (2008) plantea que esta visión es parcial y apuesta por un estudio de las políticas públicas entendidas como proceso.

Así pues, propone una reflexión sobre los aportes del análisis de las políticas públicas para entender las transformaciones de la acción pública en nuestras sociedades, a la luz de una perspectiva original fundamentada en el estudio de la acción pública como un proceso de marcos de interpretación del mundo a partir de los cuales los actores públicos y privados van a desarrollar sus estrategias para definir el contenido de las políticas públicas.

La definición y el papel de los actores en el proceso de la política reviste gran importancia para Muller, debido a que, desde su perspectiva, una política pública no es un proceso de decisión abstracto, sino que la forma de asirlo es justamente analizando las estrategias y comprendiendo, por esa vía, los comportamientos de los actores implicados.

Al respecto, el autor señala que la racionalidad de los actores no es absoluta sino parcial y fragmentada, pues no hay información perfecta que les permita analizar todos los casos posibles y además poseen distintos marcos de interpretación o maneras de ver el mundo; en consecuencia, recurren a múltiples hipótesis y las analizan de manera secuencial. De ahí, que lo que buscan sean soluciones satisfactorias y no óptimas.

Ahora bien, la importancia de la configuración de actores se sustenta en la apuesta de Muller por comprender las políticas públicas como redes de actores, cuya base se encuentra en las nociones de “redes de políticas públicas”, que pretenden entender mejor la multiplicidad de fenómenos presentes en el proceso de políticas.

Así pues, Muller sostiene que las políticas públicas “constituyen el lugar donde una sociedad construye su relación con el mundo”, por cuanto elaborar una política pública es construir una representación de la realidad sobre la cual se requiere intervenir (problema, consecuencias y soluciones posibles) Dicha representación, que no es unívoca sino múltiple, constituye el referencial de la política pública, cuya importancia radica en saber cómo esas diferentes imágenes de referencia van a encontrar un mínimo de coherencia para construir un referencial unificado. Ahí se concreta el énfasis en la función cognitiva que retoma Muller y que consiste en mostrar que la acción pública gira en torno a unos marcos propios del universo cognitivo de los actores.

En ese sentido es posible comprender el hecho de que el autor considere que las ideas son importantes para pensar las políticas públicas, lo cual explica su perspectiva cognitivista que no se reduce a la razón, pues permite conocer la configuración de relaciones con el mundo, integradas en trayectorias de dependencia y mapas mentales

que establecen concepciones sobre lo real. Además, debido a su relación con la realidad, Muller sostiene que los referenciales no son ideas pasivas o abstractas, sino ideas en acción que se juegan en espacios concretos de poder y producen significados.

Ahora bien, el proceso de elaboración de un referencial no se presenta como un proceso racional o técnico, sino que en realidad se trata de un proceso en el que operan mapas mentales a través de los cuales los actores perciben y, por esa vía, construyen la realidad. Es decir, que a partir de la imagen cognitiva o representación, los actores organizan su percepción del sistema y determinan sus comportamientos: ese conjunto de imágenes es lo que el autor denomina “referencial de la política pública” y para el que establece cuatro niveles de percepción: Por valores, lo deseable o descartable en una sociedad, lo considerado bueno o malo; por normas, lo real percibido y lo real deseado en procura de principios de acción; por algoritmos, relaciones causales que suelen ser proporcionales a los principios de acción para pensar; y por último, por imágenes, que integran los tres anteriores y consolidan el referencial.

### **1.3.2 Referenciales de la política: documentos estatales**

Habiendo esbozado de forma general el enfoque teórico del concepto de referencial de política, a partir de la obra de Pierre Muller, se presenta a continuación el desarrollo del objeto propuesto, que busca establecer cuáles son los referenciales existentes en la política pública de evaluación, desde 1991 hasta 2010 y de qué forma la política pública de evaluación en Colombia ha estado orientada a la medición de resultados.

En ese sentido integrará la identificación y análisis de los referenciales de la política pública de evaluación, a partir de diversas fuentes documentales de naturaleza estatal principalmente, que se consideran referencias obligadas para abordar la política señalada, y que serán presentados según su año de aprobación.

Así pues, los documentos estatales son tomados como dispositivos a través de los cuales operan distintas lógicas del accionar estatal “que estabilizan directrices de política, ordenándolas por temáticas que configuran el campo de controversia y estableciendo las relaciones que se dan entre ellas” (Lozano, 2008). El campo de controversia se refiere a la puesta en escena de lo que es objeto de consideración y que no ha sido resuelto, es decir que se articula a la problematización que, desde la perspectiva de Callon, adquiere una dimensión dinámica, pues trasciende la puesta en escena de una situación que se considera inadmisibile y se articula también a la creación y realización de acciones, órdenes, símbolos, etc.

De acuerdo con Lozano, “un dispositivo está constituido por conjuntos de componentes que se disponen de manera heterogénea, no operan como un sistema y no obedecen a una integración funcional” (Lozano, 2008: 82), por lo que su estudio permite identificar y analizar algunas de sus dinámicas y, desde luego, los referenciales que operan en su configuración y materialización.

La primera fuente documental a la que se recurre en materia de evaluación de políticas públicas para el periodo establecido es la Constitución Política de Colombia de 1991, que sólo menciona la evaluación en dos momentos: en el capítulo II referido a los planes de

desarrollo, en el artículo 343, atribuye a la entidad nacional de planeación el diseño y la organización de los sistemas de evaluación de gestión y resultados de la administración pública, tanto en lo relacionado con políticas como con proyectos de inversión (Constitución Política de Colombia. 1991. Capítulo II De los planes de desarrollo. Artículo 343. *Cursivas mías.*)

De esta forma, vemos que la primera mención que se hace de la evaluación, la articula a la gestión y los resultados de la administración pública, y contempla dos dimensiones, la de políticas y la de proyectos de inversión. Igualmente, se determina su organización en un sistema.

El artículo 344 del mismo capítulo, señala que los organismos departamentales de planeación harán la evaluación de gestión y resultados sobre los planes y programas de desarrollo e inversión de los departamentos y municipios (Ibíd. Artículo 344. *Cursivas mías*). Nuevamente, la evaluación se relaciona con la gestión y los resultados, pero ahora se especifica en planes y programas.

El segundo documento abordado es el decreto 2167 de Diciembre 30 de 1992<sup>2</sup> que en el capítulo I referido a estructura y funciones, artículo 2 determina entre las funciones del Departamento Nacional de Planeación “elaborar las metodologías para la preparación y evaluación de los programas y subprogramas del Plan Nacional de Desarrollo, así como para la identificación, formulación y evaluación de los proyectos financiados con los recursos del Presupuesto General de la Nación; 3. Diseñar y organizar sistemas de evaluación de gestión y de resultados de la administración, tanto en lo relacionado con políticas como con proyectos de inversión, y señalar los responsables, términos y condiciones para realizar la evaluación.

Vemos que a pesar de ser un documento expedido nueve años después de la Constitución, continúa definiendo la evaluación como articulada a programas, subprogramas y proyectos, pero ahora acota aún más al circunscribirlos al Plan Nacional de Desarrollo y a la característica de ser financiados con recursos del Presupuesto General de la Nación.

El artículo 2 del mismo documento establece las atribuciones de las dependencias en relación con el plan nacional de desarrollo y señala que las atribuciones de las dependencias que conforman la estructura del Departamento Nacional de Planeación, se ejercerán en función de las actividades de elaboración, formulación, evaluación y seguimiento del Plan Nacional de Desarrollo, así como de las políticas, estrategias, planes, programas y proyectos aprobados por el Conpes, sin perjuicio de otras disposiciones legales sobre atribución de competencias<sup>3</sup>.

Aunque aparentemente este artículo aumenta el número de elementos objeto de la evaluación, continúa acotándolos cada vez más, pues construye diversas categorías que

---

<sup>2</sup> NOTA DE VIGENCIA: Este Decreto fue derogado expresamente por el Artículo 52 del Decreto 1363 de 2000, Publicado en el Diario Oficial

<sup>3</sup> Ibíd. *Cursivas mías.*

terminan articulando la evaluación al Plan Nacional de Desarrollo, las políticas, estrategias, planes, programas y proyectos aprobados por el Conpes.

Finalmente, el artículo 22 de este decreto que habla sobre la División especial de evaluación y control de gestión, en los numerales 2 y 5, señala como sus funciones “Diseñar, homologar y difundir entre las entidades públicas, las metodologías de control de gestión y evaluación de resultados de las políticas, planes, programas y proyectos ejecutados con recursos públicos”, así como realizar la evaluación de gestión y resultados sobre los planes y programas de las entidades territoriales, que indique el Director del Departamento.

Aquí se evidencia la tipificación de la evaluación como evaluación de resultados de políticas, planes, programas y proyectos, ejecutados con recursos públicos. Esta definición implica de hecho la exclusión de otros tipos u objetivos de la evaluación y la circumscribe exclusivamente a los resultados.

Una vez trabajados los normas pertinentes para indagar acerca de los referenciales existentes en el proceso de diseño de la política pública de evaluación, se abordarán a continuación los documentos Conpes que definen los lineamientos de la política económica y social del país, y son elaborados por el Departamento Nacional de Planeación y presentados al Consejo Nacional de Política Económica y Social – CONPES-.

El Documento CONPES 2688 (DNP; 1994). Versión aprobada comienza invocando la Constitución: “La Constitución Política de 1991 ordenó la organización de un sistema de evaluación de gestión y resultados en el sector público en cabeza de la entidad nacional de planeación” y retoma los artículos 343 y 344 trabajados anteriormente.

La invocación de la Carta Política de 1991, es común a varios documentos Conpes y normas jurídicas, como elemento que legitima y justifica las acciones que planea y/o emprende el estado en el campo de la evaluación de las políticas públicas.

Más adelante, en el mismo Conpes se señala que “en 1992 se inició, bajo la dirección de la Consejería para la Modernización del Estado, un programa de mejoramiento de la gestión que en un año de labores ha involucrado a los gerentes de diversas entidades públicas en un proceso de pensamiento estratégico dirigido a promover la eficacia y eficiencia a través de la evaluación periódica de sus programas. Adicionalmente el DNP ha venido realizando evaluación ex-post de proyectos y programas seleccionados por su importante participación en los presupuestos nacionales de vigencias pasadas” (Ibíd. *Cursivas mías*).

En este documento es posible identificar de forma aún más clara el referencial que articula la evaluación a la eficacia y a la eficiencia y que define ambos criterios como ejes hacia los que debe dirigirse la política, y donde la evaluación se convierte en el medio a través del cual se pueden alcanzar. Además, señala un tipo específico de evaluación, la evaluación ex-post, de lo que se infiere que se insiste en metodologías orientadas a la medición de los resultados, es decir, durante o al finalizar la implementación de la política. Asimismo, se continúa afirmando que la evaluación se realiza sobre proyectos y programas.

Posteriormente, en este Conpes, en los principios generales se define la evaluación de resultados y el control de gestión como instrumentos de planeación que miden y promueven la calidad y oportunidad de las decisiones de los administradores públicos. Son herramientas gerenciales con capacidad de reformar y obtener una mayor efectividad en la gestión del sector público (Ibíd. *Cursivas mías*).

Se evidencia entonces, que la definición dada de evaluación de resultados se refiere a instrumentos de planeación, es decir que están diseñados para constituir un insumo en los procesos de planificación y/o formulación, que miden y promueven la calidad y oportunidad de las decisiones de los administradores públicos, nuevamente articuladas al logro de la efectividad de la gestión<sup>4</sup> del sector público. En consecuencia, la evaluación articulada a resultados, es decir que privilegia metodologías ex-post, pretende constituir un elemento útil para la planificación de las políticas; aparentemente se presenta bajo una lógica de “lección aprendida” y uno de sus objetivos más importantes es servir de insumo para la planeación y, en ciertos casos, para el rediseño (promover y reformar), aunque no se especifica de qué manera.

Este es el caso del segundo principio general, que insiste en que el énfasis del Control de Gestión o evaluación de procesos recae en el análisis de la eficiencia y eficacia de la administración en el proceso de formulación y ejecución de políticas, programas y proyectos, generando información para corregir procesos. De allí se tiene, que la evaluación genera información, que a su vez debe ser útil para corregir procesos y además, que las políticas, programas y proyectos debe sustentarse en la eficiencia y la eficacia, como se resume en el tercer principio: “La Evaluación de Resultados determina el grado de eficiencia y eficacia de la administración una vez concluidos los proyectos, programas o políticas y genera información sobre cumplimiento de metas, calidad, efectos, coberturas e impactos. Esta información provee elementos de juicio para la toma de decisiones sobre planeación y asignación de recursos”.

Finalmente, en el numeral II se define el sistema Nacional de Evaluación de Gestión y Resultados, que reitera lo señalado hasta el momento, donde la evaluación se realiza a proyectos, programas y políticas, por los Ministerios sectoriales y por el Departamento Nacional de Planeación directamente o a través de consultores privados; que las evaluaciones realizadas se complementarán entre sí sirviendo como base del análisis para formular las recomendaciones que alimentarán los subsistemas de las entidades y guiarán decisiones sobre la planificación del gasto; y que los organismos y entidades incorporarán permanentemente los resultados de las evaluaciones en los proyectos y programas en ejecución, así como en los que se encuentran en la etapa de evaluación ex-ante o de factibilidad.

El segundo documento Conpes en el que se identificarán referenciales es el 2790 (DNP; 1995), que nuevamente invoca la Constitución Política de Colombia y cuya pertinencia se explica en que constituye el documento oficial en el que se establece como directriz de la política de evaluación, la orientación a resultados.

---

<sup>4</sup> “Gestión entendida como los procesos desarrollados para facilitar la ejecución de los programas: programas entendidos como el resultado esperado de ejecución de proyectos que permiten alcanzar un objetivo de política y que se encuentra definido en el Plan de Desarrollo del gobierno” Ibíd.



Dentro de la justificación que se construye en el documento encontramos que “la estrategia para el mejoramiento de la Gestión Pública busca que la acción de las entidades del Estado gire en torno a los intereses generales de la población expresados en el Plan Nacional de Desarrollo y tenga como propósito fundamental propiciar un cambio paulatino, pero radical, en las entidades del Estado hacia una nueva cultura de la gestión pública orientada a resultados, en la cual el ciudadano sea el eje del desempeño de la administración y se persiga permanentemente el mejor aprovechamiento de los recursos”.

Allí, es posible identificar que la definición de la evaluación se articula a un lugar concreto referido a la cultura de gestión y a un problema en ella: el cambio que se requiere para que ésta esté orientada a resultados, el cual es construido y movilizado, a partir de la configuración de referenciales articulados a la medición de resultados.

Vemos que en la argumentación se afirma que los intereses de la población se encuentran expresados en el Plan Nacional de Desarrollo y a partir de eso se justifica el cambio que –se sugiere– requieren las entidades del estado, hacia una gestión pública orientada a resultados, en la que el ciudadano sea el eje de la administración y cuyo objetivo sea el mejor aprovechamiento de los recursos

Este cambio radical se sustenta en lo que se señala en el documento Conpes como el nuevo paradigma de la gestión pública<sup>5</sup>: “La gestión pública orientada a resultados” que “se basa en el fortalecimiento de dos aspectos fundamentales de la administración estatal: la gestión integral con participación ciudadana y la capacidad institucional de los organismos y entidades gubernamentales”. De esta forma, la evaluación no sólo está articulada a la medición de resultados, a la efectividad y la eficiencia, a la cultura y la gestión, sino que también es nombrada y definida a partir del lugar que configura un paradigma, el de la Nueva Gestión Pública.

Más adelante, la capacidad institucional se articulará a la eficiencia y al logro y la medición de resultados: “para que este ciclo permanente genere resultados de forma eficiente, las entidades deben contar con una capacidad institucional adecuada que les permita obtener resultados acordes con los objetivos planteados en la planeación de sus acciones. Una entidad con una adecuada capacidad institucional es aquella que cuenta con recursos suficientes y procesos de gestión adecuados para el logro de los resultados que orientan sus acciones.”

---

<sup>5</sup> La Nueva Gestión Pública –NGP– (o New Public Management NPM), es un enfoque para abordar distintos aspectos de la administración pública, que ha tenido gran incidencia desde los años 90 en Colombia, y que considera que el ciudadano en su calidad de cliente del estado puede exigir una provisión de bienes y servicios compatibles con sus preferencias y expectativas. Asimismo, sostiene que la transparencia en los sistemas presupuestales incorpora mecanismos de seguimiento, evaluación y gestión del gasto que controlan la calidad del mismo, proveen incentivos y, en teoría, aseguran disciplina fiscal. De esta forma, los sistemas de seguimiento y evaluación se encuentran contemplados en la gestión del gasto, como instrumentos para determinar la eficiencia y efectividad en provisión de bienes y servicios.

Resulta interesante observar que ahora se establecen algunas condiciones necesarias o requisitos para el logro de resultados, como la capacidad institucional, entendida como la disponibilidad de recursos suficientes y procesos de gestión acordes para lograr los resultados que se esperan de las acciones emprendidas, pues cada vez más, la evaluación, articulada a la medición de resultados, comienza a integrar diferentes aspectos de las organizaciones y su gestión.

Ahora bien, dicho cambio de paradigma que implica la transformación de las entidades se logrará mediante una estrategia propuesta en este Conpes que “tiene por objetivo general consolidar una Gestión Pública Orientada a Resultados del Plan Nacional de Desarrollo” y que se realizará, entre otros, a través de la operación del Sistema Nacional de Evaluación de Resultados de la Gestión Pública, SINERGIA, que “es una herramienta gerencial, que proveerá información pertinente y oportuna para la coordinación y evaluación inter e intrainstitucional de las acciones derivadas del Plan Nacional de Desarrollo.”

Nuevamente se define la evaluación articulada a un sistema y como herramienta gerencial cuyo resultado será información, que deberá ser útil para la coordinación de las acciones que conlleve el Plan Nacional de Desarrollo.

En el decreto 195 de enero 26 de 2004<sup>6</sup> se realizan los cambios necesarios para fortalecer la política de control de gestión y evaluación de resultados, pues se define en el artículo 2 del capítulo segundo que el Departamento Nacional de Planeación “tiene como objetivos fundamentales la preparación, el seguimiento de la ejecución y la evaluación de resultados de las políticas, planes generales, programas y proyectos del sector público y el diseño de las políticas en materia de macro estructura del Estado” y dentro de sus funciones en el artículo 4, está “aprobar las metodologías para el diseño, el seguimiento y la evaluación de las políticas, los programas y los proyectos contenidos en el Plan Nacional de Desarrollo” y “asegurar una adecuada programación del presupuesto, con base en la evaluación de resultados”.

Aunque se conserva la definición de la evaluación orientada a resultados y su ejecución sobre políticas, planes, programas y proyectos del sector público, se le atribuye la aprobación de metodologías para el diseño y la evaluación de las políticas, programas y proyectos del Plan Nacional de Desarrollo, acotando los que son susceptibles de evaluación, que ya no son todos los financiados con recursos públicos como se indicó en un primer momento, sino que ahora se limita a los contenidos en el Plan Nacional de Desarrollo. Además, se establece que existirán metodologías para la evaluación que serán definidas y aprobadas por el DNP, entidad adscrita a la rama ejecutiva del poder público, que simultáneamente es quien determina el Plan Nacional de Desarrollo; quien diseña y ejecuta las políticas es también quien las evalúa.

---

<sup>6</sup> Decreto 195 de enero 26 de 2004 *Por el cual se modifica la estructura del Departamento Nacional de Planeación.*

### 1.3.3 Conclusiones

En los procesos de diseño e implementación de la política pública de evaluación desde 1991 hasta la actualidad en Colombia, existen distintos referenciales que operan en categorías como medición de resultados, eficiencia, eficacia, evaluación ex-post, información, cultura de gestión, que definen la orientación de la política.

Tras la identificación y el análisis de los distintos referenciales operados en categorías, se concluye que en la política pública de evaluación, la medición de resultados es el referencial unificado, es decir que es el referencial construido a partir de la unión, armonía y coherencia de otros referenciales. Asimismo, se identifica como elemento común que en las trayectorias de política trabajadas, la evaluación se realiza sobre políticas, planes, programas y proyectos financiados con recursos públicos. Además, se evidencia la tipificación de la evaluación como evaluación de resultados de políticas, que implica de hecho la exclusión de otros tipos u objetivos de la evaluación y la circunscribe exclusivamente a los resultados.

De otra parte, se concluye que el texto constitucional constituye una referencia que legitima y justifica las acciones o planes del gobierno en materia de evaluación. Al ser la Constitución de 1991 el primer lugar en que el Estado se refiere a la evaluación, constituye no sólo la genealogía, sino también determina en lo fundamental las directrices de la política. De esta forma, vemos que la primera mención que se hace de la evaluación, la articula a la gestión y los resultados de la administración pública, lo cual permanecerá aún en documentos expedidos nueve años después de la Constitución.

Por otro lado, en la definición de actores involucrados en la política, el Estado y sus organismos terminan asumiendo todo el ciclo de la política pública, esto es, el diseño, la implementación, el seguimiento y la evaluación, lo cual plantea desafíos a la política de evaluación en cuanto a su objetividad y credibilidad por parte de la ciudadanía, particularmente, debido a que desde su origen, ha sido concebida como una forma de controlar la gestión y los recursos del estado, para asegurar su eficiencia y efectividad y la posibilidad de presentar resultados a la sociedad y de satisfacer sus demandas.

De otra parte, la política pública de evaluación en Colombia ha estado orientada a la medición de resultados a través de construcción de discursos, argumentos y acciones estatales que han privilegiado la identificación de información sobre los resultados o el impacto de las políticas.

De allí se tiene, que la evaluación genera información, que a su vez debe ser útil para corregir procesos, aunque no se define la forma de hacerlo claramente.

Otro elemento que debe ser considerado es que la evaluación en los dispositivos estatales analizados enfatiza en el carácter posterior de la misma, es decir, una vez iniciada su implementación o en la finalización de ésta. Sin embargo, difícilmente, una política de evaluación que se limita a la medición de resultados, deriva en un óptimo uso de la información que proporciona en los procesos de planificación y decisión de políticas.

Finalmente, se concluye que los referenciales de la política pública de evaluación en Colombia desde 1991 hasta 2010, han sido determinantes para que la orientación de dicha política corresponda principalmente a la medición de resultados.

## **2. Capítulo II: Criterios para analizar la evaluación**

En este capítulo se presentan las definiciones de los distintos criterios para evaluar o analizar una evaluación, que han llevado a establecer la ruta metodológica que se ha empleado en la presente investigación.

Este aparte resulta pertinente debido a que los conceptos tienen definiciones y acepciones diferentes que podrían conducir a aplicaciones y resultados divergentes.

### **2.1 Meta-evaluación**

Por analogía al término “evaluación” la metaevaluación (ME) es definida como la investigación sistemática realizada con el fin de hacer un juicio sobre la calidad o los méritos de una evaluación (Stufflebeam, 1981, 1987; Scriven, 1991). En palabras de Stufflebeam, la ME “es el proceso de delinear, obtener y aplicar informaciones descriptivas y de juicio –acerca de la utilidad de una evaluación, la factibilidad, pertinencia y precisión y su carácter sistemático, la competencia, la integridad/honestidad, respeto y responsabilidad social- para guiar la evaluación e informar públicamente sus fortalezas y debilidades” (2001; 183. Traducción mía). La evaluación se convierte de esta manera en el objeto de estudio de los metaevaluadores.

La especificación de la naturaleza del objeto, esto es, el tipo de evaluación que se pretende analizar, debe determinar el enfoque analítico y los métodos más apropiados para la investigación, lo cual constituye el primer paso para el metaevaluador. Así pues, el principal problema que se enfrenta cuando se pretende realizar una ME es la definición de los criterios que serán utilizados.

De hecho, este ha sido un tema recurrente entre académicos y técnicos de la ME. Sin embargo, solo hay algunas propuestas teóricas hechas específicamente para la ME; prueba de ello es que algunas de las metaevaluaciones más significativas que se han realizado hasta la fecha, corresponden a la aplicación de modelos teóricos desarrollados para las evaluaciones.

Ahora bien, las respuestas dadas por los distintos abordajes metodológicos a la pregunta sobre los criterios son diferentes aunque conservan elementos comunes.

En general, los enfoques cuantitativos tienden a enfatizar los aspectos metodológicos, particularmente, para hacer de las técnicas cuantitativas el paradigma de “calidad” de la evaluación. Por el contrario, los investigadores cualitativos parten de la consideración de la naturaleza política de la evaluación, a menudo poniendo énfasis en los procesos éticos del proceso de investigación, a veces incluso sobre las cuestiones metodológicas. Es justamente este segundo enfoque el que orienta el presente estudio.

El trabajo sobre ME de Bernstein & Freeman (1975) es un buen ejemplo de lo que es considerado “calidad de la evaluación” en investigación evaluativa, dentro del paradigma científico. De acuerdo a estos autores, el valor de una evaluación estaría garantizado solo con el uso de métodos cuantitativos para la recolección de datos como parte de un diseño de carácter experimental. En consecuencia, la finalidad de la ME consistiría en medir el grado en que la evaluación realizada se acerca a este modelo. Otros autores como Rossi (1989), Campbell y Stanley (1966) comparten esta perspectiva.

Por otro lado, a diferencia de la corriente positivista, existe la corriente naturalista en la que hay autores que han dedicado grandes esfuerzos al desarrollo teórico y a la práctica de la ME. Uno de los aportes más importantes de esta corriente es poner de relieve las deficiencias y limitaciones de los argumentos científicos positivistas, llamando la atención particularmente sobre la excesiva simplificación que conlleva equiparar la calidad de la evaluación al uso de técnicas cuantitativas.

Así, la observación de la naturaleza y las implicaciones políticas de la evaluación han llevado a autores como House (1993), Stake (1996) o Scriven (2011), entre otros, a resaltar la importancia de los aspectos éticos en el proceso de la investigación evaluativa. Estos temas se convirtieron en el estándar para medir el valor de una evaluación: en tanto la evaluación debe respetar los elementos éticos (la pluralidad de opiniones, la confidencialidad de los informantes, etc.), la ME debe centrarse también en establecer si lo ha hecho y hasta qué punto.

Scriven (2011) sin duda hace uno de los mayores aportes al campo de la ME porque va más allá de proporcionar modelos teóricos y propone criterios prácticos para la ME. El Checklist de la ME (MEC por sus siglas en inglés), propuesto por Scriven, contiene 6 elementos principales: validez, claridad, credibilidad, ética, costo-utilidad y generalización. Ahora bien, como señala el autor, el meta-evaluador puede usar una lista de estándares diseñados para la ME o puede usar una lista hecha para la evaluación.

En cualquier caso, los estándares utilizados y su justificación deben ser explícitamente detallados. El MEC debe ser susceptible de ser aplicado, preciso y riguroso (la evaluación tiene que ser rigurosa desde el punto de vista técnico, aunque se debe tener en cuenta que dar primacía al rigor puede dejar de lado aspectos tan importantes como

la relevancia de la información aportada), también debe ser sensible desde un punto de vista político, debe atender y satisfacer las necesidades de la audiencia y ser oportuna, segura, válida y "rentable" en términos de costo-efectividad.

De acuerdo con Cooksy y Caracelli (2005: 32) la metaevaluación se puede utilizar para evaluar la calidad de un solo estudio o de un conjunto de estudios. Cuando es de un único estudio la metaevaluación puede llevarse a cabo con un propósito formativo para mejorar una evaluación en curso; también puede servir a un propósito sumativo, al proporcionar información acerca de la credibilidad de los resultados de la evaluación a los usuarios y revelar las debilidades propias de los evaluadores. De esta manera, es una herramienta poderosa para aumentar la eficacia y la calidad de la práctica evaluativa.

Por otro lado, cuando la metaevaluación se centra en un conjunto de estudios tiene dos propósitos principales. En primer lugar, el conocimiento sobre la calidad de la evaluación que resulta de la metaevaluación de varias evaluaciones, se puede utilizar para informar las decisiones de los investigadores acerca de qué estudios incluir en una síntesis de evaluación<sup>7</sup>.

El segundo propósito es identificar las fortalezas y debilidades en la práctica de la evaluación con el fin de crear o fortalecer las capacidades existentes. Por ejemplo, una metaevaluación pudo encontrar que ningún estudio utilizó muestras apropiadas para responder a las preguntas de evaluación que están abordando, por lo cual sugiere la formación en técnicas de muestreo del equipo evaluador, lo que sin duda podría fortalecer las capacidades de evaluación de la agencia.

Es justamente esta perspectiva de la metaevaluación, que obedece al enfoque naturalista, la que se retoma en el presente estudio, pues lo que se busca es analizar los informes finales de las evaluaciones de impacto realizadas para identificar fortalezas y debilidades de la práctica evaluativa en Colombia, con el fin de sugerir recomendaciones que permitan mejorarla, en vez de analizar la estadística de las medidas utilizadas, así como su fiabilidad y validez a través de técnicas casi exclusivamente cuantitativas. Es por ello que el objetivo de la presente investigación no es realizar una metaevaluación de las evaluaciones de impacto, sino analizarlas a través de un instrumento diseñado especialmente para ello, con base en el abordaje metodológico de la ME y sus criterios. Esta es la razón principal para emplear el marco de análisis de la metaevaluación.

En todo caso, el objetivo final de las metaevaluaciones está orientado hacia la generación de información y análisis que permitan la creación de capacidades (*capacity building*) y por ello resultan particularmente valiosas para la comunidad de evaluación (Cooksy y Caracelli, 2005: 32 - 33).

---

<sup>7</sup> La síntesis de una metaevaluación combina la información de más de un estudio con el fin de llegar a afirmaciones generales sobre una intervención. Esta combinación puede ser cualitativa o cuantitativa. Las síntesis cualitativas han sido llamadas revisiones narrativas mientras las cuantitativas se denominan meta-análisis, que consiste en una aproximación estadística para combinar los hallazgos cuantitativos de los estudios individuales con el fin de llegar a una conclusión acerca de la eficacia general de un tipo de intervención. Las síntesis de evaluación son un enfoque común para la evaluación del impacto global de una intervención. Sin embargo, dependen de la calidad de los resultados de las evaluaciones que sintetizan (Cooksy y Caracelli, 2005: 32 - 33).

Ahora bien, como se ha señalado, la realización de una metaevaluación requiere la identificación y selección de estándares o listas de chequeo, que constituyan el referente respecto al cual se analizan las evaluaciones.

En las dos últimas décadas se han desarrollado distintos estándares en el esfuerzo por medir la calidad de las evaluaciones de impacto, entre los que se encuentran los 4 “Estándares para la Evaluación en el Sistema de Naciones Unidas” (United Nations Evaluation Group, 2005: 1-23) referidos a:

1. Marco institucional y gestión de la función de evaluación
2. Competencias y ética
3. Realización de evaluaciones
4. Informes de evaluación

También se encuentran importantes estándares para la evaluación como los *Program Evaluation Standards*, elaborados por el Joint Committee on Standards for Educational Evaluations (1994), que incluyen aspectos como la utilidad, factibilidad, propiedad y precisión de las evaluaciones y los *Estándares de Calidad para la Evaluación del Desarrollo* de la Organización de Cooperación y Desarrollo Económicos, OCDE. Los tres estándares y sus criterios serán abordados en detalle en el siguiente capítulo.

Independientemente del tipo de evaluación realizada o de los criterios usados, la evaluación debe ser diseñada para que sus resultados contribuyan a la toma de decisiones oportuna e informada respecto a los programas, a la creación de nuevas políticas o a la ampliación del conocimiento de un tema específico.

Es por ello que el diseño de la investigación –su definición, propósito y rol – es fundamental, ya que guiará todo el proceso. Desde el inicio de la evaluación, es importante tener en cuenta algunos elementos críticos en la investigación, como la perspectiva epistemológica, la selección del método más adecuado, el rigor con que debe llevarse a cabo la investigación y la validez de los resultados. Más aún, el evaluador debe ser consciente de las limitaciones de la metodología y del rigor con que se aplica. Todos estos son factores clave para que la evaluación contribuya al proceso de diseño y formulación de políticas públicas y de programas.

Adicionalmente, con base en la aproximación metodológica seleccionada para el análisis de los documentos o reportes finales de evaluación, la metaevaluación, se consideró conveniente focalizar el análisis a un tipo de evaluación para poder hacer comparables los resultados. En este caso se trabajó con las evaluaciones de impacto (EI). Esta selección, además, obedece a que la evaluación de impacto es la única de las realizadas por el estado colombiano entre 1994 y 2010, que permite establecer una relación causal entre el programa y sus efectos e impactos. Además, es la que ha requerido mayores inversiones.



## 2.2 Estándares de evaluación utilizados a nivel mundial

Las metodologías y listas de chequeo para realizar metaevaluación son una de las fuentes de información que permite construir criterios para analizar las evaluaciones. También existen distintos estándares de evaluación utilizados en todo el mundo. Están diseñados para determinar si las evaluaciones pueden ser consideradas de calidad, válidas, pertinentes y confiables. Es por ello que se ha considerado útil tener en cuenta estos estándares en el diseño del instrumento para analizar las evaluaciones de impacto (EI) realizadas por el Gobierno colombiano en el periodo de estudio.

En los últimos 20 años se han elaborado distintos estándares, particularmente en el campo de la evaluación del desarrollo, incluyendo los Estándares para la Evaluación de Programas del Joint Committee on Standards for Educational Evaluations, los estándares para la Evaluación de las Naciones Unidas (United Nations Evaluation Group, 2005), el checklist de evaluación de la Oficina de Evaluación del Ministerio Holandés para Asuntos Exteriores y los Estándares de Calidad para la Evaluación del Desarrollo del Comité de Ayuda al Desarrollo, CAD, de la Organización de Cooperación y Desarrollo Económicos, OCDE (OCDE/CAD, 2006; 2010). En este aparte se presentarán los tres principales estándares de cara a su pertinencia para el presente estudio.

Para comenzar, se encuentran los Estándares para la Evaluación de Programas del Joint Committee on Standards for Educational Evaluations (Yarbrough et al., 2011), que fueron elaborados inicialmente en 1994 y revisados recientemente. Tal como su nombre lo indica, estos criterios fueron pensados para evaluar las intervenciones realizadas en Estados Unidos en el sector educativo, particularmente en la educación básica y media. No obstante, también resultan pertinentes para la evaluación de programas y es por ello que son reconocidos no solo en la bibliografía referida a educación sino también en la de evaluación de programas.

Teniendo en cuenta los objetivos del presente estudio, se han considerado relevantes 20 (resaltados con azul) de los 30 estándares, dejando de lado 10 criterios que no son pertinentes para el objeto del presente estudio o cuya fuente de información no es el reporte final de evaluación y no fue posible acceder a ella, como es el caso de los Términos de Referencia de las evaluaciones, contenidos en el estándar de Acuerdos formales. Los 10 criterios que se dejaron de lado son: F2 Viabilidad política, U7 Impacto de la evaluación, F1 Procedimientos Prácticos, F3 Costo eficiencia, P1 Orientación de servicio, P2 Acuerdos formales, P3 Derechos de las personas, P4 Trato con las personas, A5 Validez de la información, A6 Información confiable, A7 Información sistemática.

En la siguiente tabla se presentan de manera sintética los Estándares para la Evaluación de Programas del Joint Committee on Standards for Educational Evaluations<sup>8</sup>.

**Tabla 2-1:** The Program Evaluation Standards, Joint Committee on Standards for Educational Evaluations

Program Evaluation Standards/ Estándares para la Evaluación de Programas	
Utility Standards/ Estándares de utilidad	
U1 Stakeholder Identification	U1 Identificación de actores involucrados
U2 Evaluator Credibility	U2 Credibilidad del evaluador
U3 Information Scope and Selection	U3 Alcance y selección de información
U4 Values Identification	U4 Identificación de valores
U5 Report Clarity	U5 Claridad de informes (informes eficaces)
U6 Report Timeliness and Dissemination	U6 Informes oportunos y difusión de éstos
U7 Evaluation Impact	U7 Impacto de la evaluación
Feasibility Standards/ Estándares de factibilidad	
F1 Practical Procedures	F1 Procedimientos prácticos
F2 Political Viability	F2 Viabilidad política
F3 Cost Effectiveness	F3 Costo eficiencia
Propriety Standards/ Estándares Éticos	
P1 Service Orientation	P1 Orientación de servicio
P2 Formal Agreements	P2 Acuerdos formales
P3 Rights of Human Subjects	P3 Derechos de las personas
P4 Human Interactions	P4 Trato con las personas
P5 Complete and Fair Assessment	P5 Evaluación completa y equitativa
P6 Disclosure of Findings	P6 Divulgación de los resultados (acceso)
P7 Conflict of Interest	P7 Conflicto de intereses
P8 Fiscal Responsibility	P8 Responsabilidad fiscal
Accuracy Standards/ Estándares de precisión	
A1 Program Documentation	A1 Documentación del Programa

## <sup>8</sup> En el Anexos

Anexo A: Joint Committee on Standards for Educational Evaluations , se encuentra una definición en español detallada para cada uno de los criterios.

Program Evaluation Standards/ Estándares para la Evaluación de Programas	
A2 Context Analysis	A2 Análisis de contexto
A3 Described Purposes and Procedures	A3 Procesos y procedimientos documentados
A4 Defensible Information Sources	A4 Fuentes de información defendibles
A5 Valid Information	A5 Validez de la información
A6 Reliable Information	A6 Información confiable
A7 Systematic Information	A7 Información sistemática
A8 Analysis of Quantitative Information	A8 Análisis de la información cuantitativa
A9 Analysis of Qualitative Information	A9 Análisis de la información cualitativa
A10 Justified Conclusions	A10 Conclusiones justificadas
A11 Impartial Reporting	A11 Reporte imparcial
A12 Metaevaluation	A12 Metaevaluación

Fuente: Elaboración propia a partir de *The Program Evaluation Standards*, Joint Committee on Standards for Educational Evaluations, 1994 (Traducción mía)

Como se observa, se han dejado por fuera aquellos criterios referidos a la validez o calidad de la información, puesto que, como se ha señalado, el objetivo de este análisis no es hacer una ME ni retomar el enfoque positivista de la ME para analizar la estadística de las medidas utilizadas en las evaluaciones.

Cabe señalar que uno de los aportes de esta propuesta que se ha considerado importante, es que se incluye la metaevaluación como uno de los estándares de calidad de la evaluación.

En segundo lugar, se encuentran los estándares de evaluación del Sistema de Naciones Unidas. Estos estándares tienen un perfil mucho más orientado a la evaluación de programas para el desarrollo, con enfoque de derechos humanos y de género y se realizan con el ánimo de lograr efectividad en las intervenciones de desarrollo y para rendir cuentas a los gobiernos o autoridades locales así como a los donantes internacionales.

Muchos de los criterios propuestos coinciden con los estándares del Joint Committee on Standards for Educational Evaluations, presentado previamente, por lo que a continuación se han incluido solamente aquellos que se consideran pertinentes para la presente investigación (en azul), y que complementan los criterios del Joint Committee seleccionados. En total se han considerado pertinentes 8 estándares de la metodología del Sistema de Naciones Unidas<sup>9</sup>.

De la misma manera que con los estándares del Joint Committee, se han dejado de lado aquellos criterios referidos a la validez de la información, puesto que, como se ha señalado, el objetivo de este análisis no es hacer una ME ni medir la calidad de la estadística utilizada en las evaluaciones.

A continuación se presenta la Tabla 2-2 que contiene los Estándares de evaluación en el Sistema de Naciones Unidas.

---

<sup>9</sup> En el Anexo B: Estándares de evaluación en el Sistema de las Naciones Unidas, se encuentran los estándares completos.

**Tabla 2-2** Estándares de evaluación en el Sistema de las Naciones Unidas

Estándares de evaluación en el Sistema de las Naciones Unidas	
1. Marco institucional y gestión de la función de evaluación	
Marco institucional	
Estándar 1.1: Las organizaciones de las NU deben contar con un marco institucional adecuado para la gestión efectiva de su función de evaluación.	<p>Promover una cultura que valore la evaluación como una base para el aprendizaje.</p> <p>Facilitar un proceso de evaluación independiente e imparcial, garantizando que la función de evaluación se realice de manera independiente de otras funciones gerenciales. El jefe/a de la unidad de evaluación deberá reportar directamente al órgano de gobierno de la organización o al jefe/a de ésta.</p>
Estándar 1.2: Las organizaciones de las NU deben elaborar una política de evaluación y actualizarla regularmente, teniendo en consideración las Normas y los Estándares de Evaluación del sistema de las NU.	
Estándar 1.3: Las organizaciones de las NU deben asegurar que los planes de evaluación de actividades de evaluación sean presentados a sus órganos de gobierno y/o a los jefe/as de organizaciones para su revisión y/o aprobación.	
Estándar 1.4: Las organizaciones de las NU deben garantizar la existencia de mecanismos apropiados para el seguimiento de las evaluaciones; y formular una política explícita en materia de divulgación.	6. Una política de divulgación debe garantizar la difusión transparente de los resultados de las evaluaciones, lo cual incluye poner los informes a disposición general de los órganos de gobierno y del público, salvo en aquellos casos en que se requiera protección y confidencialidad razonables para algunas de las partes interesadas.
Gestión de la función de evaluación	
Estándar 1.5: El jefe/a de evaluación ejerce un rol de liderazgo en garantizar que la función de evaluación se encuentre plenamente operativa y que el trabajo en materia de evaluación, se lleve a cabo de acuerdo con los más altos estándares profesionales.	
Estándar 1.7: El jefe/a de evaluación debe cerciorarse de que la función de evaluación sea dinámica, adaptándose a los nuevos desarrollos y a las necesidades cambiantes tanto dentro como fuera de la organización.	
2. Competencias y ética	1. Todas las personas involucradas en el diseño, la realización y la gestión de actividades de evaluación deben aspirar a realizar un trabajo de alta calidad, que se rija por estándares profesionales y principios éticos y morales.
3. Realización de evaluaciones	

Estándares de evaluación en el Sistema de las Naciones Unidas	
Estándar 3.7: Las metodologías de evaluación deben ser lo suficientemente rigurosas como para evaluar la intervención y garantizar una evaluación exhaustiva, justa y libre de sesgos.	12. Los métodos de evaluación dependen de la información que se busque y del tipo de datos que se esté analizando. Los datos deben provenir de una variedad de fuentes a fin de garantizar la exactitud, validez y confiabilidad de los mismos, y que todas las personas afectadas y/o partes interesadas sean tomadas en cuenta. La metodología debe incidir explícitamente en los aspectos de género y de grupos subrepresentados.
	13. Deben reconocerse también, las limitaciones de los métodos de evaluación que fueron seleccionados.
Estándar 3.9: El diseño de la evaluación debe incluir, donde sea relevante, consideraciones sobre hasta qué punto el compromiso del sistema de las NU con el enfoque basado en los derechos humanos ha sido incorporado en el diseño de la intervención que será evaluada.	20. El diseño de la evaluación podría incluir además, algún proceso de revisión ética del diseño original de la intervención que será evaluada, que incluya lo siguiente:
	un balance de costos y beneficios para los participantes, incluyendo el potencial impacto negativo; las consideraciones éticas que subyacen a quién se incluye o excluye de la evaluación y cómo se lleva a cabo este proceso;
	el manejo de la privacidad y la confidencialidad; las prácticas para obtener acuerdos informados; la retroalimentación a los participantes; los mecanismos para determinar y monitorear la conducta y la práctica de los evaluadores/as y recolectores/as de datos.
Proceso	
Estándar 3.10: La relación entre el evaluador/a y la institución o instituciones que comisionan la evaluación tiene que caracterizarse, desde el principio, por el respeto mutuo y la confianza.	
Estándar 3.15: Las evaluaciones deben llevarse a cabo de una manera profesional y ética.	34. Las evaluaciones deben llevarse a cabo de una manera integral y balanceada, de manera tal que las diferentes perspectivas sean consideradas y analizadas. Los hallazgos clave deben ser sustentados por medio de la triangulación. Cualquier conflicto de interés debe ser abordado de una manera abierta y honesta, de modo tal que no socave el resultado de la evaluación.
	35. Los evaluadores/as deben discutir, de una manera apropiada para el contexto, aquellos valores, supuestos, teorías, métodos, resultados y análisis que pudieran afectar significativamente la interpretación de los hallazgos de la evaluación. Esto se aplica a todos los aspectos de la evaluación, desde su concepción inicial hasta el uso eventual de los hallazgos.
Informes	
Estándar 3.16: El informe final de evaluación debe contar con una estructura lógica, contener hallazgos basados en evidencia, conclusiones, lecciones y recomendaciones y estar libre de información irrelevante para el análisis general. El informe debe ser presentado en un formato que permita que la	37. El lector o lectora del informe de evaluación debe poder entender:
	el propósito de la evaluación; qué, exactamente, se evaluó; cómo se diseñó y llevó a cabo la evaluación; qué evidencias se encontró; qué conclusiones se extrajeron; qué recomendaciones se hicieron; qué lecciones se derivaron.

Estándares de evaluación en el Sistema de las Naciones Unidas	
información sea comprensible y accesible.	39. Los evaluadores/as deben permitir a todos los actores pertinentes, acceso a información procedente de la evaluación que sea apropiada, y, de ser posible, deben difundir activamente dicha información entre las partes interesadas. Las comunicaciones a una parte interesada deben incluir, siempre, los resultados importantes que pudieran afectar sus intereses. En todos los casos, los evaluadores/as deben esforzarse por presentar los resultados con la mayor claridad y sencillez posible, de manera tal que los clientes y otras partes interesadas puedan entender fácilmente el proceso y los resultados de la evaluación.
Seguimiento	
Estándar 3.17: La evaluación requiere de una respuesta explícita por parte de las autoridades de gobierno y gerencia afectadas por sus recomendaciones.	
4. Informes de evaluación	
Estándar 4.2: El informe de evaluación debe contener un resumen ejecutivo.	
Estándar 4.10: La evaluación debe proporcionar una descripción detallada de la participación de las partes interesadas.	
Estándar 4.12: En la presentación de los hallazgos es necesario, en la medida de lo posible, medir los insumos, productos y efectos / impactos (o dar una explicación apropiada de por qué no se hace).	20. Los efectos e impactos deben incluir cualquier resultado(s) no intencional(es), ya sean beneficiosos o perjudiciales. En adición, deberá incluirse cualquier efecto(s) multiplicador(es) o a nivel micro, de la intervención que está siendo evaluada. En la medida de lo posible, cada uno de éstos efectos deberá ser medido ya sea cuantitativa o cualitativamente. Al utilizar dichas mediciones, deberá hacerse referencia a parámetros referenciales.
Estándar 4.14: Deben ser identificadas, en la medida de lo posible, las razones de los logros y las dificultades de la intervención que está siendo evaluada, especialmente los factores de restricción y habilitación.	26. El informe de evaluación debe ir más allá de una simple descripción de la implementación y de los efectos e incluir un análisis, basado en los hallazgos, de las causas subyacentes, las restricciones, las fortalezas que hay que capitalizar y las oportunidades. Deben identificarse y analizarse, en la medida de lo posible, los factores externos que contribuyen a los logros y dificultades, entre ellos la situación social, política o ambiental.

Fuente: Elaboración propia a partir de *Estándares de Evaluación en el Sistema de las Naciones Unidas*, Grupo de Evaluación de las Naciones Unidas (UNEG), 2005.

Finalmente, se encuentran los Estándares de Calidad para la Evaluación del Desarrollo del Comité de Ayuda al Desarrollo, CAD, de la Organización de Cooperación y Desarrollo Económicos, OCDE, que fueron diseñados en 2006 con el objeto de establecer un parámetro o referente para la calidad de los procesos y productos de la evaluación en los países miembros, justamente para mejorar la efectividad de las intervenciones orientadas al desarrollo de los países. Por lo tanto, tienen en cuenta como un actor relevante en las evaluaciones a los Organismos Internacionales como donantes de recursos o asistencia técnica a quienes debe rendirse cuentas.

Los estándares pretenden servir como una guía para mejorar las evaluaciones de las intervenciones de desarrollo, que son fundamentales para determinar su impacto real. Adicionalmente, los estándares fueron diseñados para contribuir a la armonización de los esfuerzos de evaluación de los distintos países miembros. Específicamente, los estándares tienen como objetivo:

- Proporcionar normas para el proceso y los productos (outputs) de las evaluaciones;
- Facilitar la comparación de las evaluaciones de los distintos países;
- Facilitar las alianzas y la colaboración en las evaluaciones conjuntas;
- Permitir que los países miembros utilicen los resultados y los informes de evaluación de los demás (incluidas las buenas prácticas y las lecciones aprendidas);
- Agilizar los esfuerzos de evaluación (CAD/OCDE, 2006: 3)

Inicialmente, los Estándares tenían diez componentes principales, aprobados por los miembros de la Red de Evaluación del CAD en su cuarta reunión del 30 y 31 de marzo de 2006 para su aplicación en fase de prueba durante tres años, a saber: 1. Motivo, propósito y objetivos de una evaluación; 2. Alcance de la evaluación; 3. Contexto; 4. Metodología de evaluación; 5. Fuentes de información; 6. Independencia; 7. Ética de la evaluación; 8. Calidad de la evaluación; 9. Relevancia de los resultados de la evaluación; y 10. Exhaustividad (CAD/OCDE, 2006). La lista completa de los estándares y sus descripciones se presentan en el Anexo C.

Sin embargo, tras la realización de una encuesta en 2008 sobre el uso de los Estándares y un seminario celebrado en 2009 en Auckland, que se sumaron a los comentarios recibidos por los miembros de la Red del DAC de Evaluación del Desarrollo, se logró consolidar una versión final de los estándares en 2010 (CAD/OCDE, 2010).

En vez de los primeros diez componentes, los Estándares de Calidad para la Evaluación de Desarrollo actualmente solo tienen cuatro componentes principales: 1. Consideraciones generales; 2. Finalidad, planificación y diseño; 3. Ejecución y presentación de resultados; y 4. Seguimiento, utilización y aprendizajes. Mientras varios estándares fueron incluidos en uno de los cuatro componentes principales, algunos fueron incorporados en otros estándares<sup>10</sup>.

Uno de los valores agregados de la propuesta del CAD/OCDE se encuentra en que son estándares contruidos mediante un consenso internacional, que es la naturaleza de esta organización.

A continuación, en la Tabla 2-3 se presenta una síntesis de estos Estándares. Los que se han considerado pertinentes para esta investigación se encuentran resaltados en azul.

---

<sup>10</sup> Por ejemplo, la Evaluación de resultados y el muestreo que se incluyeron previamente bajo el componente de la Metodología de Evaluación, se incluyeron en la Explicación de la metodología utilizada y la Validez y fiabilidad de las fuentes de información, respectivamente. Por otra parte, la Validez y fiabilidad de las fuentes de información incluye los dos estándares relacionados con las fuentes de información (Transparencia de las fuentes de información y Fiabilidad y exactitud de las fuentes de información).

**Tabla 2-3** Estándares de Calidad para la Evaluación del Desarrollo de la OCDE-CAD 2010

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2010	
1. Consideraciones generales	
1.1. Evaluación del desarrollo	Por evaluación se entiende la determinación del valor o importancia de una intervención de esta naturaleza.
1.2. Un proceso evaluativo libre y abierto	Para una mayor credibilidad, el proceso evaluación ha de ser transparente e independiente de la gestión programática o la elaboración de políticas en la materia.
1.3. Ética de la evaluación	Los evaluadores respetan las principales normas y códigos de conducta profesionales y éticos que les atañen, así como los derechos humanos y las diferencias culturales y en las costumbres, creencias y prácticas religiosas de todos los actores implicados. También son sensibles a las consideraciones de género, origen étnico, capacidad, edad, orientación sexual, lengua y otras diferencias.
1.4. Enfoque asociativo	El concepto de asociación lleva implícita la idea de proceso incluyente en el que participan diversos actores interesados, como los poderes ejecutivo y legislativo, la sociedad civil, los supuestos beneficiarios y los socios internacionales.
1.5. Coordinación y alineamiento	El proceso de evaluación ha de tener en cuenta los planes, actividades y políticas nacionales y locales en la materia.
1.6. Desarrollo de capacidad	Deben maximizarse los efectos positivos del proceso de evaluación sobre la capacidad evaluativa de los socios para el desarrollo.
1.7. Control de calidad	El control de calidad se ejerce a lo largo de todo el proceso de evaluación.
2. Finalidad, planificación y diseño	
2.1. Fundamentación y finalidad de la evaluación	La fundamentación, finalidad y utilización prevista de la evaluación deben indicarse claramente, mencionándose por qué se efectúa la evaluación en ese momento concreto, para quién se lleva a cabo y cómo se emplearán sus resultados a efectos de aprendizaje y rendición de cuentas.
2.2. Objetivos específicos de la evaluación	Los objetivos específicos de la evaluación deben establecer claramente el propósito de la misma.
2.4. Evaluabilidad	Debe determinarse si se ha definido adecuadamente la intervención, si sus resultados son verificables y si la evaluación constituye la mejor respuesta a las cuestiones planteadas por los actores implicados relevantes o por los responsables de las políticas.
2.5. Participación de los actores relevantes	Los actores implicados relevantes deben intervenir en las primeras fases del proceso y debe dárseles la oportunidad de contribuir a su diseño, incluyendo la identificación de los asuntos que han de abordarse y las preguntas a las que la evaluación ha de responder



Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2010	
2.6. Consideración sistemática de la evaluación conjunta	Entre organismos donantes y los países socios.
2.7. Cuestiones a las que pretende responder la evaluación	Preguntas evaluativas específicas y relevantes
2.8. Selección y aplicación de los criterios de evaluación	En la evaluación se aplican los criterios adoptados por el CAD: pertinencia, eficiencia, eficacia, impacto y sostenibilidad.
2.9. Selección del enfoque y de la metodología	La metodología se elabora de acuerdo con el enfoque elegido y debe incluir la especificación y justificación del diseño de la evaluación y las técnicas de recogida y análisis de datos. La metodología seleccionada debe responder a las preguntas de la evaluación utilizando evidencias creíbles.
2.10. Recursos	Los recursos disponibles para la evaluación deben ser suficientes en cuanto a fondos, personal y habilidades técnicas, de manera que los objetivos de la misma puedan cumplirse de forma eficaz.
2.11. Estructuras de gobierno y gestión	Las estructuras de gobierno y gestión se deben diseñar para adaptarse al contexto, propósito, alcance y objetivos de la evaluación.
	La estructura de gobierno debe salvaguardar la credibilidad, la transparencia y el carácter integrador de la evaluación.
2.12. Documentación relativa a la finalidad y las expectativas	"Términos de Referencia" (TdR)
3. Ejecución y presentación de resultados	
3.1. Equipo de la evaluación	Debe seguirse un procedimiento de contratación abierto y transparente para la selección del equipo evaluador.
3.2. Independencia de los evaluadores frente a los interesados	Los evaluadores deben ser independientes respecto de la intervención para el desarrollo de que se trate, incluido de sus funciones de formulación de políticas, operativas y de gestión, así como sus beneficiarios previstos.
3.3. Consulta y protección de los implicados	Se debe proteger los derechos y el bienestar de los participantes en la evaluación.
3.4. Realización de la evaluación dentro del plazo y presupuesto asignados	La evaluación debe realizarse y sus resultados facilitarse a quienes la encomendaron en la forma oportuna y que permita alcanzar los objetivos previstos. La evaluación debe llevarse a cabo de forma eficiente y dentro del presupuesto asignado. Las modificaciones de las condiciones y circunstancias han de ponerse de manifiesto, y se explicarán, tratarán y consensuarán entre las partes los cambios imprevistos de plazo y presupuesto.
3.5. Informe de la evaluación	El informe de la evaluación ha de ser fácilmente comprensible por sus destinatarios, y su modalidad será la adecuada a la vista de la finalidad del ejercicio. El informe ha de abarcar los siguientes elementos y apartados: (numerales 3.6 a 3.15)
3.6. Claridad y representatividad del resumen	
3.7. Contexto de la intervención para el desarrollo	

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2010	
3.8. Lógica de la intervención	
3.9. Validez y fiabilidad de las fuentes de información	
3.10. Explicación de la metodología utilizada	En el informe se debe describir y explicar la metodología evaluativa y su aplicación. Para analizar efectos directos e impactos, se explicará la atribución y/o contribución de los resultados. En el informe se reconocerán las posibles restricciones halladas y su efecto sobre la evaluación, incluyendo sobre su independencia e imparcialidad. Se han de detallar las técnicas empleadas para la recogida y análisis de los datos. Se justificarán las opciones y se explicarán las limitaciones y deficiencias.
3.11. Claridad del análisis	
3.12. Respuestas a las preguntas evaluativas	
3.13. Mención de los cambios y las limitaciones de la evaluación	
3.14. Mención de las discrepancias entre los miembros del equipo	
3.15. Inclusión de las observaciones de los actores implicados	Se debe ofrecer a los actores implicados relevantes la posibilidad de formular observaciones al borrador de informe. En el texto definitivo se deben reflejar dichas observaciones y poner de manifiesto las discrepancias sustanciales.
4. Seguimiento, utilización y aprendizajes	
4.1. Oportunidad, pertinencia y utilización de la evaluación	Se garantiza una difusión sistemática y la conservación y gestión del informe de manera que todos los socios para el desarrollo puedan acceder fácilmente al mismo, se llegue a los destinatarios previstos y se aprovechen al máximo los beneficios de la evaluación.
4.2. Respuesta sistemática a las recomendaciones y seguimiento de las mismas	Las personas/entidades destinatarias de las recomendaciones deben responder de forma sistemática a las mismas y adoptar las medidas procedentes. Esto incluye una respuesta y un seguimiento formales por parte de la dirección. Las medidas acordadas han de ser supervisadas para garantizar que se rinde cuentas de su aplicación.
4.3. Difusión	Los resultados de la evaluación se presentan en un formato accesible y se difunden de forma sistemática, tanto interna como externamente, para extraer lecciones aprendidas, realizar acciones de seguimiento y garantizar la transparencia. A la luz de las enseñanzas surgidas del ejercicio, se señalan otros posibles interesados dentro de la amplia comunidad del desarrollo y se entabla contacto con ellos de forma que se saque el máximo partido.

Fuente: Elaboración propia con base en: Organización para la Cooperación y el Desarrollo Económicos –OCDE-, Comité de Ayuda al Desarrollo –CAD-, Estándares de Calidad para la Evaluación del Desarrollo, 2010.

Aparentemente, los criterios referidos a la metodología reciben menos atención en la nueva versión de los estándares del CAD/OCDE, en comparación con la versión de 2006. Esto resulta sorprendente ya que, en comparación con otros estándares de evaluación utilizados con frecuencia, las normas de evaluación del CAD/OCDE 2006 ya hacían menos énfasis que las demás en las cuestiones metodológicas. Por el contrario, en la tercera edición de los Estándares de Evaluación Educativa de programas del Joint Committee incluso se refuerza la atención en la metodología y las fuentes de información con dos estándares que cubren este tema (información válida y fiable y gestión de la información).

Los estándares difieren respecto al nivel de detalle pero en general capturan aspectos como la validez, la confiabilidad, la ética y la utilidad de las evaluaciones. En el siguiente acápite se presentará el análisis de los tres estándares con ocasión del diseño del instrumento que se empleará para el estudio de los informes finales de las EI realizadas por el gobierno colombiano entre 1994 y 2010.

## **2.3 Metodología: Instrumento diseñado para el análisis de las evaluaciones estatales de impacto**

Para comenzar, es importante señalar que de acuerdo al objetivo principal de esta investigación que consiste en conocer y analizar la configuración y operación de la política pública de evaluación en Colombia, desde 1994 y hasta 2010, para identificar sus fundamentos y definir sus alcances y limitaciones, con el fin de sugerir recomendaciones que permitan enriquecerla o transformarla, se contempla el análisis crítico de los informes finales de las evaluaciones de impacto realizadas por el Estado colombiano en el periodo de estudio. De tal forma, se ha considerado que el enfoque metodológico de la ME y los estándares definidos para la evaluación que existen actualmente, resultan pertinentes para el abordaje de las evaluaciones y para su análisis, debido a que incluyen preguntas y etapas útiles que permiten orientar la sistematización y el análisis de la información más relevante de los informes finales de evaluación, por lo que el instrumento metodológico diseñado incluye un conjunto de criterios contruidos con base en dichos estándares.

En primer lugar, retomando los estándares del Sistema de Naciones Unidas, se ha construido un estándar denominado Marco institucional, que indaga sobre la función evaluación en el arreglo institucional correspondiente. Su objetivo es identificar si el proceso de *evaluación ha sido independiente e imparcial, garantizando que la función evaluación se realice de manera independiente de otras funciones gerenciales* y velando por que los posibles conflictos de interés sean presentados y gestionados de manera abierta.

Este criterio se complementa con el estándar referido a Ejecución y presentación de resultados del DAC/OCDE (2010), en tanto indaga sobre la independencia de los evaluadores frente a los interesados señalando que *los evaluadores deben ser independientes respecto de la intervención para el desarrollo de que se trate, incluido de sus funciones de formulación de políticas, operativas y de gestión, así como sus beneficiarios previstos*.

También del DAC/OCDE (2010) se retoma para este estándar el criterio de Consideraciones generales referido al Enfoque asociativo que *lleva implícita la idea de proceso incluyente en el que participan diversos actores interesados, como los poderes ejecutivo y legislativo, la sociedad civil, los supuestos beneficiarios y los socios internacionales.*

Finalmente, en este estándar se ha incluido el criterio referido al Conflicto de intereses de los estándares del Sistema de Naciones Unidas y del Joint Committee on Standards for Educational Evaluations, que comparten este criterio, y que señala que estos *deben ser presentados de manera abierta y honesta para que no afecten la evaluación y sus resultados.*

El segundo estándar se denomina Enfoque y metodología y es el que será útil para analizar la aproximación metodológica de los informes finales de evaluación. Para la elaboración de este estándar se retoman los aportes del DAC/OCDE (2010), del Joint Committee y del Sistema de Naciones Unidas, e incluye tres criterios: Identificación de valores, perspectivas, procedimientos y lógicas utilizadas para interpretar los hallazgos de la evaluación que deben ser cuidadosamente detallados. En ese sentido, los evaluadores/as deben discutir, de una manera apropiada para el contexto, aquellos valores, supuestos, teorías, métodos, resultados y análisis que pudieran afectar significativamente la interpretación de los hallazgos de la evaluación. Esto se aplica a todos los aspectos de la evaluación, desde su concepción inicial hasta el uso eventual de los hallazgos. Este criterio tiene dos sub-criterios: el primero denominado Selección del enfoque y de la metodología y que pretende verificar si la metodología se elabora de acuerdo con el enfoque elegido y debe incluir la especificación y justificación del diseño de la evaluación y las técnicas de recogida y análisis de datos. El segundo sub-criterio indaga por las limitaciones de los métodos de evaluación que fueron seleccionados, así como por los cambios y las limitaciones de la evaluación y las discrepancias entre los miembros del equipo.

El tercer estándar retoma aportes de los tres estándares referidos y está orientado a determinar la calidad del informe final de la evaluación, que debe ir más allá de una simple descripción de la implementación y de los efectos e incluir un análisis, basado en los hallazgos, de las causas subyacentes, las restricciones, las fortalezas que hay que capitalizar y las oportunidades.

Este estándar se compone de cinco sub-criterios: i) Conclusiones justificadas (las conclusiones de la evaluación deben estar explícitamente justificadas y responder a los hallazgos y análisis realizados); ii) Imparcialidad (el reporte debe evitar sesgos y distorsiones causadas por intereses o creencias personales, para que refleje exactamente los hallazgos); iii) Claridad (en todos los casos, los evaluadores/as deben esforzarse por presentar los resultados con la mayor claridad y sencillez posible, de manera tal que los clientes y otras partes interesadas puedan entender fácilmente el proceso y los resultados de la evaluación. Los lectores del informe de evaluación debe poder entender: el propósito de la evaluación; qué, exactamente, se evaluó; cómo se diseñó y llevó a cabo la evaluación; qué evidencias se encontró; qué conclusiones se extrajeron; qué recomendaciones se hicieron; qué lecciones se derivaron); e iv) Impactos no intencionales (los efectos e impactos deben incluir cualquier resultado(s) no intencional(es), ya sean beneficiosos o perjudiciales. En la medida de lo posible, cada

uno de estos efectos deberá ser medido ya sea cuantitativa o cualitativamente); v) Factores externos y de contexto (deben identificarse y analizarse, en la medida de lo posible, los factores externos que contribuyen a los logros y dificultades, entre ellos la situación social, política o ambiental)

El cuarto estándar se refiere al Impacto de la evaluación e indaga por la oportunidad, pertinencia y utilización de la evaluación. El estándar se compone de 3 sub-criterios: el primero consiste en garantizar una difusión transparente y sistemática de los resultados de las evaluaciones para que todos los actores interesados (gobierno, organismos internacionales y sociedad civil) puedan acceder fácilmente al mismo, se llegue a los destinatarios previstos y se aprovechen al máximo los beneficios de la evaluación. Este sub-criterio ha sido construido con base en el estándar del Joint Committee, en el que la divulgación de resultados se refiere más a asegurar que todos los hallazgos sean accesibles para todas las personas afectadas por la evaluación y para aquellos que tengan derecho legal a conocer los resultados, similar a la de los estándares del Sistema de Naciones Unidas en el que la difusión está orientada hacia la disposición general de los órganos de gobierno y del público de los informes de evaluación. También se retoma la propuesta DAC/OCDE en la que se habla de una difusión sistemática orientada a sacar el mayor provecho posible de los resultados de la evaluación.

El segundo sub-criterio se refiere al uso de la evaluación para la toma de decisiones. El tercer sub-criterio se refiere a la Metaevaluación y está orientado a determinar si *la evaluación en sí misma ha sido evaluada usando como base estándares pertinentes que permitan examinar sus fortalezas y debilidades*, la intención es analizar las razones por las cuales esas evaluaciones han resultado más o menos influyentes y, de ser posible, las vías mediante las cuales han producido su influencia. Este sub-criterio ha sido tomado de la propuesta del Joint Committee, que es el único de los estándares que la incluye y se considera de vital importancia, por cuanto la metaevaluación debe ser una práctica propia de los sistemas de evaluación de políticas públicas, como herramienta de control de calidad, de aprendizaje y de mejoramiento de la práctica evaluativa.

A continuación se presenta el instrumento diseñado con base en los estándares existentes de evaluación y de ME.

**Tabla 2-4** Estándares para el análisis de las evaluaciones de impacto conducidas por el DNP en Colombia, 1994 - 2010

Estándares para el análisis de las evaluaciones de impacto conducidas por el DNP en Colombia, 1994 - 2010	
1. Marco institucional	
1.1 El proceso de evaluación ha sido independiente e imparcial	La función de evaluación se realiza de manera independiente de otras funciones gerenciales
	Los evaluadores deben ser independientes respecto de la intervención para el desarrollo de que se trate, incluido de sus funciones de formulación de políticas, operativas y de gestión, así como sus beneficiarios previstos.

Estándares para el análisis de las evaluaciones de impacto conducidas por el DNP en Colombia, 1994 - 2010	
1.2 Enfoque asociativo	La evaluación es el resultado de un proceso incluyente en el que participan diversos actores interesados, como los poderes ejecutivo y legislativo, la sociedad civil, los supuestos beneficiarios y los socios internacionales.
1.3 Conflicto de intereses	Los posible conflictos de intereses deben ser presentados de manera abierta y honesta para que no afecten la evaluación y sus resultados
2. Enfoque y metodología	
2.2 Selección del enfoque y de la metodología	La metodología se elabora de acuerdo con el enfoque elegido y debe incluir la especificación y justificación del diseño de la evaluación y las técnicas de recogida y análisis de datos
2.3 Limitaciones de los métodos de evaluación	Se deben referir las limitaciones de los métodos de evaluación que fueron seleccionados, así como los cambios y las limitaciones de la evaluación y las discrepancias entre los miembros del equipo sobre las metodologías aplicadas en caso de existir.
3. Informe final	
3.1 Conclusiones justificadas	Las conclusiones de la evaluación deben estar explícitamente justificadas y responder a los hallazgos y análisis realizados.
3.2 Imparcialidad	El reporte debe evitar sesgos y distorsiones causadas por intereses o creencias personales, para que refleje exactamente los hallazgos
3.3 Claridad	Los resultados deben presentarse con la mayor claridad y sencillez posible, de manera tal que los clientes y otras partes interesadas puedan entender fácilmente el proceso y los resultados de la evaluación: el propósito de la evaluación; qué, exactamente, se evaluó; cómo se diseñó y llevó a cabo la evaluación; qué evidencias se encontró; qué conclusiones se extrajeron; qué recomendaciones se hicieron; qué lecciones se derivaron
3.4 Impactos no intencionales	Los efectos e impactos deben incluir cualquier resultado(s) no intencional(es), ya sean beneficiosos o perjudiciales. En la medida de lo posible, cada uno de estos efectos deberá ser medido ya sea cuantitativa o cualitativamente
3.5 Factores externos y de contexto	Deben identificarse y analizarse, en la medida de lo posible, los factores externos que contribuyen a los logros y dificultades, entre ellos la situación social, política o ambiental
4. Impacto de la evaluación	
4.1 Difusión transparente y sistemática de los resultados de la evaluación	Garantizar una difusión transparente y sistemática de los resultados de las evaluaciones para que todos los actores interesados
4.2 Uso de la evaluación	La evaluación debe ser útil para la toma de decisiones

Estándares para el análisis de las evaluaciones de impacto conducidas por el DNP en Colombia, 1994 - 2010	
4.3 Metaevaluación	La evaluación en sí misma debe ser evaluada usando como base estándares pertinentes que permitan examinar sus fortalezas y debilidades.

Fuente: Elaboración propia.

Por último, cabe señalar que para la conformación de la propuesta metodológica de análisis de la evaluación estatal de programas –objetivo de esta tesis –, de este capítulo ha sido importante: a) establecer los distintos criterios que existen para evaluar o analizar evaluaciones; b) caracterizar las posibilidades y limitaciones de los mismos; y c) definir un instrumento que resulte pertinente para el caso colombiano y permita responder a las preguntas de investigación. La identificación de todos estos elementos permite tener claridad sobre el tipo de componentes que habrían de tenerse en cuenta a la hora de analizar los informes finales de evaluación.

## 2.4 Selección de sistema de calificación y definición de la muestra

Una vez presentado el instrumento metodológico diseñado para hacer el análisis de las evaluaciones de impacto realizadas por el gobierno colombiano, este aparte detalla la muestra y el sistema de calificación definidos.

Es importante señalar que, tal como fue mencionado en el contexto histórico, a pesar de la existencia de SINERGIA desde el año 1994, es solo hasta 2002 cuando el gobierno decide dar un impulso importante al Sistema e iniciar evaluaciones de políticas públicas. Por lo tanto, en 2006 se contrata la primera evaluación de impacto.

Las evaluaciones de impacto están orientadas a establecer los efectos generados en la población beneficiaria después de haber recibido la intervención de un programa o política social. Determina cuál es el efecto en las condiciones de los beneficiarios debido a la exposición al programa, aislando otros posibles factores que incidan en esta población. Por lo tanto, permite tomar decisiones de política en la focalización del gasto social que soportan el proceso de rediseño, mejora o eliminación de los programas si se estima necesario. Como valor agregado, el conocimiento adquirido permitirá diseñar adecuadamente programas y proyectos futuros.

Entre 2006 y 2010 se realizaron en total nueve (9) evaluaciones de impacto que se relacionan en la **Tabla 2-4**.

**Tabla 2-4:** Evaluaciones de impacto estatales finalizadas de 1994 a 2010

Nombre de la Evaluación	Tipo	Sector	Evaluador	Año
Indicadores de goce efectivo de derechos para población desplazada	Impacto	Presidencia	UT Econometría - SEI	2010

Nombre de la Evaluación	Tipo	Sector	Evaluador	Año
La formación del proceso de formación musical y el impacto sobre el desempeño académico del Proyecto Preorquestal de BATUTA	Impacto	Presidencia	UT Economía Urbana - SEI	2010
Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)	Impacto	Ambiente y Vivienda	Centro de Estudios Regionales, Cafeteros y Empresariales (CRECE)	2009
Evaluación de Impacto del programa Hogares Comunitarios de Bienestar - HCB	Impacto	Salud y Protección	UT Universidad de los Andes - Profamilia	2008
Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento	Impacto	Presidencia	Centro Nacional de Consultoría	2008
Evaluación del Programa Jóvenes en Acción	Impacto	Presidencia	UT Econometría - Fedesarrollo - SEI - IFS	2007
Evaluación de impacto del Programa de Vivienda de Interés Social Urbana -VISU (Intermedia 1483)	Impacto	Ambiente y Vivienda	Centro de Estudios Regionales, Cafeteros y Empresariales (CRECE)	2007
Programa Familias en Acción: Impactos en capital humano - Evaluación rural	Impacto	Presidencia	UT Econometría-IFS-SEI	2007
Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias	Impacto	Salud y Protección	UT Econometría-IFS-SEI	2006

Fuente: SINERGIA. Evaluaciones finalizadas (DNP) [Última consulta: 29 Octubre 2013]. Disponible en: <https://sinergia.dnp.gov.co/PortalDNP/default.aspx>

De acuerdo a la metodología de la metaevaluación usada internacionalmente, ésta se realiza sobre evaluaciones de un mismo programa o de programas con objetivos e instrumentos similares. Sin embargo, como el objetivo del presente estudio no es realizar una metaevaluación sino analizar desde una perspectiva crítica el conjunto de evaluaciones de impacto realizadas por el DNP entre 1994 y 2010, se abordarán a través del instrumento creado las 9 evaluaciones independientemente de que sus temáticas correspondan a distintos objetivos de política.

El universo muestral consiste en las 9 evaluaciones y este análisis se basa en una muestra de 5 evaluaciones de impacto que cubren el periodo de estudio (1994-2010), aunque como se ha señalado las evaluaciones de impacto en el país se empezaron a realizar efectivamente en el 2004. Los criterios para seleccionar esta muestra fueron los siguientes: Para evaluar los cambios en el transcurso del tiempo, se incluyó en la muestra 1 programa que fue evaluado dos veces mediante evaluaciones de impacto (Familias en Acción). Aparte de este criterio, y el de asegurar que exista una evaluación por año, en la selección de los programas no se dio preferencia a ningún sector. En consecuencia, la muestra para el análisis incluye las siguientes evaluaciones de impacto.



**Tabla 2-5** Muestra seleccionada para el análisis

Nombre de la Evaluación	Tipo	Sector	Evaluador	Año
E1 Indicadores de goce efectivo de derechos para población desplazada	Impacto	Presidencia	UT Econometría - SEI	2010
E2 Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)	Impacto	Ambiente y Vivienda	Centro de Estudios Regionales, Cafeteros y Empresariales (CRECE)	2009
E3 Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento	Impacto	Presidencia	Centro Nacional de Consultoría	2008
E4 Programa Familias en Acción: Impactos en capital humano - Evaluación rural	Impacto	Presidencia	UT Econometría-IFS-SEI	2007
E5 Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias	Impacto	Salud y Protección	UT Econometría-IFS-SEI	2006

Fuente: Elaboración propia

### 2.4.1 Sistema de calificación<sup>11</sup>

El análisis de los informes finales de evaluación se realizó de acuerdo a los criterios construidos en el instrumento diseñado para ello, con base en la aplicación de estándares internacionales, y que han sido explicados en el aparte anterior: Marco institucional (El proceso de evaluación ha sido independiente e imparcial, Enfoque asociativo, Conflicto de intereses), Enfoque y metodología (Selección y Limitaciones), Informe final (Conclusiones justificadas, Imparcialidad, Claridad, Impactos no intencionales, Factores externos y de contexto) e Impacto de la evaluación (Difusión transparente y sistemática de los resultados de la evaluación, Uso de la evaluación, Metaevaluación).

Cada uno de los 9 informes ha sido analizado en función de estos criterios y sub-criterios, y la calificación resultante es producto de un riguroso análisis de los informes finales de evaluación que constituyen la principal fuente de información para esta parte del estudio.

La escala de calificación definida,:

---

<sup>11</sup> La metodología propuesta en el documento del Banco Mundial titulado “Chile: Análisis del Programa de Evaluación del Gasto Público” (Rojas, et al., 2005) aportó importantes luces en la elaboración del instrumento así como en el sistema de calificación.

1) Altamente Satisfactoria. Puntajes de 91 a 100 indican que la evaluación realizada cumple con todos los criterios establecidos y que hace aportes importantes para la gestión de los programas gubernamentales.

2) Satisfactoria. Los puntajes de 76 a 90 indican que la evaluación cumple con la mayoría de los criterios y presenta algunos aportes para la gestión de los programas pero que éstos no tienen la relevancia del nivel anterior.

3) Moderadamente Satisfactoria. Los puntajes de 50 a 75, indican que la evaluación realizada cumple con la mayoría de los criterios de calidad y presenta algunos aportes.

4) Insatisfactoria. Los puntajes inferiores a 49 indican que la evaluación no ha cumplido con la mayoría de los criterios de calidad establecidos y no presenta aportes significativos.

Para ubicar cada uno de los informes en esta escala de calificación es necesario definir pesos para cada uno de los estándares, que han sido asignados de acuerdo a la importancia del criterio o sub-criterio, en el marco del objetivo del estudio. La suma de los pesos asignados a cada nivel es 100%. La escala de calificación definida, no sólo permite calificar y ubicar el estudio en un nivel de calidad, sino además organizar las evaluaciones realizadas para el análisis de conjunto

**Tabla 2-5** Escala de calificación

DETALLE	POR SUB-CRITERIO			POR CRITERIO	
	Nota	Peso	Puntaje	Peso	Puntaje
<b>1. MARCO INSTITUCIONAL</b>	<b>0</b>		<b>0</b>	20%	0
1.1 El proceso de evaluación ha sido independiente e imparcial		30%	0		
1.2 Enfoque asociativo		30%			
1.3 Conflicto de intereses		40%			
<b>2. ENFOQUE Y METODOLOGÍA</b>	<b>0</b>		<b>0</b>	25%	0
2.1 Selección del enfoque y de la metodología		50%	0		
2.2 Limitaciones de los métodos de evaluación		50%	0		
<b>3. INFORME FINAL</b>	<b>0</b>		<b>0</b>	30%	0
3.1 Conclusiones justificadas		25%			
3.2 Imparcialidad		25%			
3.3 Claridad		20%			
3.4 Impactos no intencionales		10%			
3.5 Factores externos y de contexto		20%			
<b>4. IMPACTO DE LA EVALUACIÓN</b>	<b>0</b>		<b>0</b>	25%	0
4.1 Difusión transparente y sistemática de los resultados de la evaluación		25%	0		
4.2 Uso de la evaluación		35%	0		
4.3 Metaevaluación		40%	0		

<b>PUNTAJE GENERAL</b>				100%	0
------------------------	--	--	--	------	---

Fuente: Elaboración propia.

La justificación de la ponderación de los pesos de cada sub-criterio se presenta en la siguiente tabla.

**Tabla 2-6** Justificación de la asignación de los pesos

Criterios	Peso asignado	Justificación
1. Marco institucional	20%	La independencia e imparcialidad de la evaluación se han considerado medianamente importantes en el marco de este análisis, aunque reciben la menor ponderación debido a que el énfasis del análisis se centrará en el informe final, la metodología y el impacto de la evaluación.
2. Enfoque y metodología	25%	Teniendo en consideración que el abordaje propuesto en el presente estudio hacia la metodología utilizada en las evaluaciones, no pretende determinar la validez ni rigurosidad de la estadística, se le ha asignado un menor peso que al informe final. Así pues aunque el peso asignado a este criterio no es el mayor, tiene la misma ponderación que el criterio referido al impacto de la evaluación.
3. Informe final	30%	En el contexto global del presente estudio y teniendo en cuenta que los informes finales de evaluación son la fuente principal de información para el análisis, se le ha asignado el mayor peso de todos los criterios.
4. Impacto de la evaluación	25%	Debido a que este criterio indaga por la difusión, utilización y metaevaluación de la evaluación que resultan fundamentales para comprender esta parte del proceso de la práctica evaluativa en Colombia, se le ha otorgado la misma ponderación que al enfoque y la metodología.

Fuente: Elaboración propia.

Para poder ponderar la calificación obtenida por cada uno de los informes y de acuerdo a la metodología sugerida por Stufflebeam (2001: 199 - 200), es necesario establecer el grado en que el informe de evaluación cumple o no cada criterio y sub-criterio, de forma que se pueda asignar una nota tras el análisis del informe, que permita, con base en el peso asignado, determinar el puntaje alcanzado. A continuación se presentan los criterios para establecer el grado de cumplimiento de los criterios y sub-criterios.

**Tabla 2-7** Calificación del grado de cumplimiento de los criterios y sub-criterios de cada estándar

Estándar	Sub-criterio	Grado de cumplimiento de criterios		
		Cumplido 80 - 100	Parcialmente cumplido 60 - 79	No cumplido <60

Estándar	Sub-criterio	Grado de cumplimiento de criterios		
		Cumplido 80 - 100	Parcialmente cumplido 60 - 79	No cumplido <60
1. Marco institucional	1.1 El proceso de evaluación ha sido independiente e imparcial	Se cumplen los 3 sub-criterios o 2 se cumplen y al menos 1 se cumple parcialmente	Al menos 2 sub-criterios se cumplen totalmente	Ningún sub-criterio se cumple
	1.2 Enfoque asociativo			
	1.3 Conflicto de intereses			
2. Enfoque y metodología	2.2 Selección del enfoque y de la metodología	Se cumplen los 2 sub-criterios	Al menos 1 sub-criterio se cumple total o parcialmente	Ningún sub-criterio se cumple
	2.3 Limitaciones de los métodos de evaluación			
3. Informe final	3.1 Conclusiones justificadas	Se cumplen los 5 sub-criterios o 4 se cumplen y al menos 1 se cumple parcialmente	Por lo menos 4 de los 5 sub-criterios se cumplen total o parcialmente	Tres o más sub-criterios no se cumplen.
	3.2 Imparcialidad			
	3.3 Claridad			
	3.4 Impactos no intencionales			
	3.5 Factores externos y de contexto			
4. Impacto de la evaluación	4.1 Difusión transparente y sistemática de los resultados de la evaluación	Se cumplen los 3 sub-criterios o 2 se cumplen y al menos 1 se cumple parcialmente	Por lo menos 2 de los 3 sub-criterios se cumplen total o parcialmente	Tres o más sub-criterios no se cumplen.
	4.2 Uso de la evaluación			
	4.3 Metaevaluación			

Fuente: Elaboración propia.

Así pues, de acuerdo con esta escala y con base en los estándares contruidos se calcula un índice para cada criterio que al final da una puntuación al informe de evaluación, con el fin de identificar las fortalezas y debilidades de la evaluación.

Finalmente, se da una clasificación general para la evaluación que se retoma del Checklist del OCDE/CAD (2006), que no sólo permite calificar y ubicar el estudio en un nivel de calidad, sino además organizar las evaluaciones realizadas para el análisis de conjunto. Esta escala contempla 4 niveles de satisfacción con el cumplimiento de los criterios de cada evaluación:

Débil = 2

Parcialmente satisfactorio = 3

Satisfactorio = 4

Excelente = 5

### 3. Capítulo III: ¿Qué nos dicen las evaluaciones? Presentación de hallazgos y análisis.

Una vez presentado el instrumento metodológico y el sistema de calificación diseñados, así como la muestra definida para hacer el análisis de las evaluaciones de impacto realizadas por el gobierno colombiano, este capítulo incluye los principales hallazgos y el análisis derivado de éstos.

Este aparte pretende responder a la segunda pregunta de investigación: ¿Cuáles son las principales características de la evaluación que realiza el Estado colombiano a través del Departamento Nacional de Planeación?

Para realizar el análisis de los informes de las Evaluaciones de Impacto, se aplicaron los siguientes criterios construidos en el instrumento diseñado para ello, con base a la aplicación de estándares internacionales: Marco institucional (El proceso de evaluación ha sido independiente e imparcial, Enfoque asociativo, Conflicto de intereses), Enfoque y metodología (Selección y Limitaciones), Informe final (Conclusiones justificadas, Imparcialidad, Claridad, Impactos no intencionales, Factores externos y de contexto) e Impacto de la evaluación (Difusión transparente y sistemática de los resultados de la evaluación, Uso de la evaluación, Metaevaluación).

La Tabla 3-1 ofrece una visión sintética de los resultados cuantitativos, mientras que las secciones 3.1. a 3.4. discuten brevemente el cumplimiento de los informes de las 5 evaluaciones respecto a cada estándar, destacando tendencias o casos significativos. El detalle de la calificación obtenida por cada una de las evaluaciones se encuentra en el Anexo D.

**Tabla 3-1** Puntajes de los 5 Informes finales de las evaluación de impacto

DETALLE	Peso	EI 1	EI 2	EI 3	EI 4	EI 5
		Puntaje	Puntaje	Puntaje	Puntaje	Puntaje
<b>1. MARCO INSTITUCIONAL</b>	20%	1,4	2,8	1,4	1,4	1,4
1.1 El proceso de evaluación ha sido independiente e imparcial						
1.2 Enfoque asociativo						
1.3 Conflicto de intereses						

DETALLE	Peso	EI 1	EI 2	EI 3	EI 4	EI 5
		Puntaje	Puntaje	Puntaje	Puntaje	Puntaje
<b>2. ENFOQUE Y METODOLOGÍA</b>	25%	8,8	5,3	0,0	4,4	4,4
2.1 Selección del enfoque y de la metodología						
2.2 Limitaciones de los métodos de evaluación						
<b>3. INFORME FINAL</b>	30%	3,9	3,3	2,6	2,6	2,0
3.1 Conclusiones justificadas						
3.2 Imparcialidad						
3.3 Claridad						
3.4 Impactos no intencionales						
3.5 Factores externos y de contexto						
<b>4. IMPACTO DE LA EVALUACIÓN</b>	25%	1,5	1,8	2,0	2,0	0,0
4.1 Difusión transparente y sistemática de los resultados de la evaluación						
4.2 Uso de la evaluación						
4.3 Metaevaluación						
<b>PUNTAJE GENERAL</b>	100%	3,9	3,3	1,5	2,6	1,9

Fuente: Elaboración propia

Como se puede observar, de acuerdo a la escala de análisis retomada del Checklist del OCDE/CAD (2006): Débil = 2, Parcialmente satisfactorio = 3, Satisfactorio = 4 y Excelente = 5, ninguno de los informes de evaluación tiene un desempeño excelente y solo dos de ellos se acercan a una calificación satisfactoria. Por el contrario 3 de los 5 estudios se consideran débiles de acuerdo a los criterios y sub-criterios establecidos para el análisis.

A continuación se presentan los hallazgos para cada uno de los criterios y sub-criterios ponderados, así como su análisis.

### 3.1 Marco institucional

La mayoría de las evaluaciones contratadas por el DNP (3 de las 5 evaluaciones analizadas) fueron con la misma Unión Temporal, conformada por el Institute of Fiscal Studies, Econometría y Sistemas Especializados de Información S.A.

La única firma de toda la muestra que no es un consultor de evaluación común o de los contratados usualmente es el Centro de Estudios Regionales, Cafeteros y Empresariales (CRECE), que estuvo a cargo de la evaluación del Programa de Vivienda de Interés Social Urbana –VISU.

La evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento, fue contratada con el Centro Nacional de Consultoría, que si bien solo

es el evaluador de este programa en la muestra definida, ha participado en al menos 8 evaluaciones de las contratadas por el DNP.

Para el caso de Colombia esto resulta particularmente importante debido a que el mercado de consultoría para evaluar políticas es pequeño y excluyente en razón, entre otros, a los activos financieros que se exigen como respaldo en las licitaciones de las evaluaciones de impacto.

Solamente en la evaluación del Programa de Vivienda de Interés Social Urbana –VISU, se cumple en sub-criterio de Enfoque asociativo, pues se tuvieron en cuenta actores distintos a los beneficiarios, que tenían relación con la implementación del programa como Constructores, Cajas de Compensación Familiar y Entes Territoriales, a través de entrevistas semiestructuradas y grupos focales. En los demás estudios, la evaluación se realiza con enfoque *top-down*, de arriba hacia abajo que no evidencia la participación de distintos actores, aparte del gobierno, la firma contratada para realizar la evaluación y los beneficiarios que, dicho sea de paso, tienen un rol pasivo (informantes) y no participan del proceso de evaluación.

Para ninguno de los 5 estudios fue posible encontrar evidencia de conflictos de intereses surgidos entre los actores durante la evaluación. Sin embargo, esto no permite concluir que no se hayan presentado, sino que es posible que los informes finales de evaluación no sean la fuente de información adecuada para medir este sub-criterio. De existir esta información, podría estar consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.

### 3.2 Enfoque y metodología

Todas las evaluaciones usaron la misma metodología cuasi-experimental que consiste en la comparación entre grupo beneficiario y grupo de control, para establecer la situación con y sin programa. Este diseño metodológico corresponde a la perspectiva científica tradicional y coincide con la evaluación de tercera generación planteada por Guba y Lincoln (1989).

En la evaluación del Programa de Vivienda de Interés Social Urbana, aunque se utilizó la metodología cuasi experimental, se complementó con las metodologías cualitativas (grupos focales). Lo mismo ocurrió en ambas evaluaciones del Programa Familias en Acción, en las que se incluyeron diálogos colectivos e individuales con diversos tipos de actores del programa.

Particular atención en este estándar merece el caso de la Evaluación del Programa Empleo en Acción, en el que se tuvo que cambiar el diseño experimental por uno cuasi-experimental, debido a que el grupo de control no había sido elegido aleatoriamente.

Adicionalmente, existían importantes diferencias en la Línea de Base entre la población de personas seleccionadas inicialmente como beneficiarios y los participantes que finalmente trabajaron en las obras, debido a las demoras en su iniciación. En consecuencia, el tamaño del grupo de control terminó siendo la mitad del grupo de tratamiento, y además se tuvo que excluir un municipio, San Vicente del Caguán, como

un caso extremo porque se alejó de la forma en la cual debía operar el Programa. También debido a la demora, 47 de los 155 proyectos que se encuestaron en línea de base fueron eliminados porque no se realizaron.

De tal forma, en esta evaluación no se consideraron los problemas en el inicio y la operación del programa en los distintos territorios por lo que la metodología tuvo que ser ajustada, comprometiendo la validez de los resultados.

El sub-criterio referido a las imitaciones de la metodología solo fue ponderado positivamente en el caso de la evaluación de Indicadores de goce efectivo de derechos para población desplazada, porque en la evaluación se reconocen algunas de las limitaciones de la metodología: por ejemplo que algunos indicadores cambiaron en la forma de medición respecto a la línea base, por lo que no son comparables con los datos de 2008, y se consideran por lo tanto una nueva línea base. Adicionalmente, se señala que la concentración por razones de eficiencia de la selección de la muestra de hogares en las cabeceras municipales del país y área periféricas, impide conocer las situación de los desplazados que se han ubicado en zonas rurales o que han retornado a su sitio de residencia en dicha zona, agravada por el hecho de que no se puede estimar que proporción de dicha población ha elegido este tipo de lugar de residencia.

En la evaluación del Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias, aunque no se identifican las limitaciones de la metodología, se señalan las dificultades enfrentadas durante el proceso de recolección de información que comprometen los hallazgos y resultados de la evaluación.

### 3.3 Informe final

Para los cinco informes finales de evaluación analizados, las conclusiones se justifican con base en los datos y análisis contenidos en el informe. Sin embargo, en el caso de la evaluación de Indicadores de goce efectivo de derechos para población desplazada, no se puede perder de vista el problema de comparabilidad que existe entre los datos de tratamiento y la línea base.

Algo similar ocurre con el Programa Empleo en Acción en el que los ajustes metodológicos realizados debido a las dificultades presentadas, comprometen la confiabilidad de las conclusiones aunque estas respondan al diseño metodológico que finalmente se adoptó.

Adicionalmente, la mayoría de las evaluaciones se concentran exclusivamente en los objetivos explícitos del programa y el uso de las metodologías cuantitativas limita la posibilidad de medir (o al menos observar) otros hallazgos. Por ejemplo, en el caso de la evaluación del goce efectivo de derechos de la población desplazada, el fallo de la Corte Constitucional y los parámetros que ordenó medir, representaron un desafío para la evaluación que, en general se considera que no fue respondido adecuadamente.

En este caso, los evaluadores consideraron que el indicador de resultado del goce efectivo de derechos, GED (propuesto por el gobierno nacional y complementado por la Corte Constitucional) a una vivienda digna, es *muy exigente y sigue siendo bajo*.



Claramente este es un sesgo de los evaluadores, que al medir el logro de este indicador coinciden en que es bajo y pretenden explicar (¿excusar?) este resultado señalando que el indicador es muy exigente. En las demás evaluaciones no se registra evidencia de sesgos en el proceso de evaluación.

En general, las evaluaciones no contemplan impactos no intencionales. La excepción se da en el marco del Programa de Vivienda de Interés Social Urbana, que aunque no registra impactos no intencionales, gracias a la aplicación de metodologías cualitativas, establece que la consecución de la vivienda podría tener un efecto negativo en los procesos de organización y participación comunitaria, y formula la hipótesis según la cual es posible que al lograr el objetivo que unía a los beneficiarios los lazos se debilitaron.

Los factores externos y de contexto no son tenidos en cuenta en ninguna de las evaluaciones analizadas. Normalmente, lo que se encuentra es una breve descripción del programa, pero el contexto de la evaluación y del programa no se tienen en cuenta.

Por ejemplo, en el caso de la evaluación del programa Familias en Acción, hubo algunas dificultades durante la primera evaluación, debido al cambio de gobierno. Sin embargo este tema no se menciona en el informe final de evaluación. Briceño et al (2011: 9, Traducción mía) describe lo ocurrido en los siguientes términos: “De acuerdo con los actores involucrados que fueron entrevistados para este estudio, la negociación de los Términos de Referencia (TdR) y del contrato de la evaluación entre la oficina de Presidencia y el Equipo de la Dirección de Evaluación de Políticas Públicas del DNP, fue muy complicada, al punto que los gerentes de la evaluación de impacto creyeron que el proceso no iba a prosperar. Por un lado, no se existía un precedente de una evaluación de este tipo y alcance en el país, y sobre todo, algunos funcionarios del gobierno consideraban que los costos eran escandalosos. La negociación empezó en octubre del 2000 y el contrato fue firmado finalmente a finales de diciembre de 2001”.

Llama la atención que esta situación no sea mencionada en el informe final de la evaluación, aun cuando en este caso el contexto política afectaba la realización de la evaluación. Al respecto, cabe señalar que esta es una de las debilidades de Sinergia y particularmente de la práctica evaluativa en Colombia en el periodo estudiado: los procesos no son documentados en detalle, pues las evaluaciones enfatizan en el componente técnico cuantitativo de la evaluación, dejando de lado otros aspectos relevantes como el contexto (político, social, institucional, etc.).

Otro buen ejemplo de la importancia del contexto en el proceso de evaluación ocurre durante la primera evaluación de impacto de Familias en Acción, cuando el equipo evaluador tuvo que enfrentar un cambio inesperado, como consecuencia de la decisión del gobierno colombiano de expandir el programa tan pronto como fuera posible, por cuanto las elecciones presidenciales estaban próximas.

Briceño et al (2006: 10, Traducción mía) explican la situación en los siguientes términos: “El equipo de evaluación planeaban realizar el levantamiento de la línea base entre junio y octubre de 2002. Sorpresivamente, a comienzos de 2002, la Oficina de Presidencia le informó al equipo que ellos empezarían a desembolsar los pagos en todos los municipios de tratamiento inmediatamente, evitando la posibilidad de levantar una línea de base pura.”

En este caso, nuevamente el contexto político opera como un factor externo que afecta la evaluación, pero tampoco es mencionado en el informe final de evaluación, sino que simplemente se explica el ajuste metodológico requerido para mantener la confiabilidad y la validez de la línea base.

### **3.4 Impacto de la evaluación**

En general en ninguno de los documentos se encuentra información referida a si los resultados de la evaluación fueron utilizados para la toma de decisiones sobre el programa o para el control ciudadano. En algunos casos y como parte de la revisión bibliográfica realizada durante la investigación fue posible encontrar evidencias de la no utilización de las evaluaciones, en otras fuentes documentales.

El caso de la evaluación de Empleo en Acción, lo que resulta aún más preocupante es el informe final se publicó en el año 2006 y el programa se había suspendido en 2003, lo que indica que perdió toda pertinencia y la posibilidad de ser utilizada se redujo significativamente.

En el caso de Familias en Acción, si bien en el reporte final de evaluación no se registra evidencia de su uso para la toma de decisiones, el Gobierno Nacional tomó inicialmente la decisión de prorrogar su duración hasta diciembre de 2006 y posteriormente se diseñó una segunda fase de expansión del Programa para el periodo 2007- 2010, tal como se recomendó en las evaluaciones.

Adicionalmente, aportó elementos para ajustar el programa orientados a la reducción del subsidio para escolaridad primaria dados los altos niveles de cobertura, a la introducción de incrementos graduales en el monto de subsidios de secundaria y a la creación de premios por graduación de secundaria.

Cabe señalar que tras el ejercicio realizado, se ha concluido que los informes finales de evaluación no resultan ser la fuente de información idónea para indagar sobre el impacto de la evaluación, su uso ni sobre la difusión del mismo.

En ese sentido los informes parecen estar orientados a cumplir eficazmente con la noción ordinaria de la notificación pública la cual, como sustituto de la rendición de cuentas, se limita a la presentación de resultados, a través de los medios institucionales establecidos como la aplicación web de Sinergia Evaluaciones (disponible a través de la página web de Sinergia) y el Boletín de Evaluaciones de Política Pública.

Finalmente, cabe señalar que ninguna de las evaluaciones analizadas incluyó la metaevaluación posterior, lo que representa un desafío importante y una tarea pendiente en la práctica evaluativa en Colombia, pues la metaevaluación es un ejercicio que aporta información significativa para mejorar la realización de evaluaciones.

## **4. Conclusiones y recomendaciones**

### **4.1 Conclusiones**

Aún cuando las evaluaciones de impacto realizadas por el DNP y, especialmente, las del programa Familias en Acción, frecuentemente son destacadas como un caso relevante para los expertos en evaluación de impacto, todavía hay un margen importante para mejorar, especialmente de cara a la metaevaluación, la metodología, el contexto y el uso de las evaluaciones.

En términos de la independencia e imparcialidad, aún cuando el criterio fue calificado de buena forma, especial atención merece el reducido mercado de evaluadores que existe en Colombia, así como el hecho de que solo un grupo de firmas sea las que realiza la mayoría de las evaluaciones de impacto en el país.

Respecto al uso de la evaluación, en los 5 casos analizados, no se evidencia que los resultados de la evaluación hayan incidido en el proceso de toma de decisiones frente al programa evaluado. En algunos casos, como el de Empleo en Acción, los resultados de la evaluación no resultaron pertinentes pues fueron presentados después de que la decisión sobre el programa ya había sido tomada.

Aunque se conservan en la actualidad muchos de los elementos de la visión clásica de evaluación, se insiste en la necesidad de que la evaluación de las políticas públicas trascienda su sentido instrumental como herramienta de medición, y avance hacia la construcción de nuevas ideas, metodologías y herramientas.

### **4.2 Recomendaciones**

Después de realizar este análisis de la práctica evaluativa en Colombia, con énfasis en las evaluaciones de impacto y basada en las conclusiones formuladas, se sugieren 4

recomendaciones para mejorar la práctica evaluativa en general en el país, y específicamente la de las evaluaciones de impacto.

La importancia del contexto sociopolítico y económico de la evaluación debe ser reconocido y valorado en la realización de evaluaciones, porque este puede afectar el proceso de la evaluación (lo que en efecto ocurrió en el caso de Familias en Acción).

Las habilidades y conocimientos en el tema de evaluación en el país, fuera de las firmas consultoras y al interior de la administración pública deben ser promovidos, pues esto contribuye a generar capacidad institucional.

El uso de la evaluación debe ser fortalecido y fomentado desde el gobierno nacional, pues aún resulta incierto para los procesos de toma de decisión y para el control ciudadano. Tal vez se debería diseñar, implementar o, si ya existe, transformar las estrategias de difusión de los resultados de los informes y acercarlos a los espacios de debate y control ciudadanos en un lenguaje comprensible y adaptado a todo tipo de público.

Por lo tanto, se recomienda explorar nuevas formas de difusión y de promoción del conocimiento y uso de las evaluaciones por parte de la ciudadanía, a través de mecanismos como alianzas con organizaciones de la sociedad civil, una estrategia de comunicaciones que tenga en cuenta los públicos, los contextos y los medios más adecuados para comunicar de una forma simple y asertiva a ciudadanos (no especializados en el tema) los principales hallazgos de las evaluaciones y sus implicaciones para la administración pública y en la toma de decisiones.

Finalmente, todo proceso de evaluación constituye una experiencia enriquecedora, por lo que la sistematización del aprendizaje que implicó su realización puede constituirse en guía de futuras evaluaciones y mejoras permanentes de la metodología aplicada. En tal sentido, se recomienda que los evaluadores incluyan un acápite específico en que describan las Lecciones Aprendidas, tanto las referentes a la ejecución del programa, como las del proceso de evaluación.

## Anexos

### A. Anexo A: Joint Committee on Standards for Educational Evaluations

Joint Committee on Standards for Educational Evaluations (1994)		
Program Evaluation Standards/ Estándares para la Evaluación de programas		
Utility Standards/ Estándares de utilidad		Definición
U1 Stakeholder Identification	U1 Identificación de actores involucrados	Actores involucrados en la evaluación y actores afectados por la evaluación.
U2 Evaluator Credibility	U2 Credibilidad del evaluador	Confiable, competente.
U3 Information Scope and Selection	U3 Alcance y selección de información	Responder a preguntas pertinentes e incluir necesidades e intereses del cliente o de otros actores involucrados.
U4 Values Identification	U4 Identificación de valores	Perspectivas, procedimientos y lógicas utilizadas para interpretar los hallazgos de la evaluación deben ser cuidadosamente detallados (son la base del juicio de validez).
U5 Report Clarity	U5 Claridad de informes (informes eficaces)	La información esencial del programa evaluado así como los hallazgos de la evaluación deben ser incluidos y presentados en una forma de fácil comprensión.

Joint Committee on Standards for Educational Evaluations (1994)		
Program Evaluation Standards/ Estándares para la Evaluación de programas		
U6 Report Timeliness and Dissemination	U6 Informes oportunos y difusión de éstos	Los informes de evaluación y las conclusiones provisionales significativas deben ser difundidos a los usuarios previstos para que puedan ser utilizados en el momento oportuno.
U7 Evaluation Impact	U7 Evaluación de impacto	La forma en que se realiza la evaluación deberá promover el seguimiento por parte del responsable que permita aumentar las probabilidades de que la evaluación sea utilizada.
Feasibility Standards/ Estándares de factibilidad		Definición
F1 Practical Procedures	F1 Procedimientos prácticos	Reducen las interrupciones al mínimo mientras se levanta la información.
F2 Political Viability	F2 Viabilidad política	Anticipar las posiciones de grupos de interés para lograr su cooperación. Permitiría reducir o contrarrestar posibles obstáculos y sesgos en la evaluación.
F3 Cost Effectiveness	F3 Costo eficiencia	La evaluación debe ser eficiente y producir información válida que justifique el uso de los recursos disponibles.
Propriety Standards/ Estándares éticos		Definición
P1 Service Orientation	P1 Orientación de servicio	Evaluaciones que atiendan y respondan a las necesidades de la población objetivo.
P2 Formal Agreements	P2 Acuerdos formales	Términos de referencia de la evaluación claros y escritos.
P3 Rights of Human Subjects	P3 Derechos de las personas	La evaluación debe respetar y proteger los derechos y el bienestar de las personas.
P4 Human Interactions	P4 Trato con las personas	Los evaluadores deben respetar a las personas con las que interactúan.
P5 Complete and Fair Assessment	P5 Evaluación completa y equitativa	La revisión y el registro de las fortalezas y debilidades del programa debe ser completa y equitativa para que sea útil.

Joint Committee on Standards for Educational Evaluations (1994)		
Program Evaluation Standards/ Estándares para la Evaluación de programas		
P6 Disclosure of Findings	P6 Divulgación de los resultados	Las partes involucradas en la evaluación deben asegurar que todos los hallazgos sean accesibles para todas las personas afectadas por la evaluación y para aquellos que tengan derecho legal a conocer los resultados.
P7 Conflict of Interest	P7 Conflicto de intereses	Los conflictos de intereses deben ser presentados de manera abierta y honesta para que no afecten la evaluación y sus resultados.
P8 Fiscal Responsibility	P8 Responsabilidad fiscal	La asignación de recursos y los gastos del evaluador deben reflejar procesos de rendición de cuentas, ser prudentes y su uso ético y responsable.
Accuracy Standards/ Estándares de precisión		Definición
A1 Program Documentation	A1 Documentación del Programa	El programa evaluado debe ser descrito y documentado de forma clara y precisa.
A2 Context Analysis	A2 Análisis de contexto	El contexto en el cual el programa existe debe ser examinado en detalle para identificar sus efectos e influencia en el programa.
A3 Described Purposes and Procedures	A3 Procesos y procedimientos documentados	Los procesos y procedimientos de la evaluación deben ser monitoreados y descritos detalladamente.
A4 Defensible Information Sources	A4 Fuentes de información defendibles	Las fuentes deben ser descritas en detalle para lograr la exactitud de la información.
A5 Valid Information	A5 Validez de la información	Los procedimientos de levantamiento de información deben ser diseñados e implementados para asegurar que los análisis que surjan sean válidos para el uso previsto de la evaluación.
A6 Reliable Information	A6 Información confiable	La información es suficiente confiable para el uso de la evaluación.
A7 Systematic Information	A7 Información sistemática	La información en una evaluación debe ser revisada sistemáticamente.

Joint Committee on Standards for Educational Evaluations (1994)		
Program Evaluation Standards/ Estándares para la Evaluación de programas		
A8 Analysis of Quantitative Information	A8 Análisis de la información cuantitativa	La información debe ser analizada de manera apropiada y sistemáticamente para responder a las preguntas de la evaluación.
A9 Analysis of Qualitative Information	A9 Análisis de la información cualitativa	
A10 Justified Conclusions	A10 Conclusiones justificadas	Las conclusiones deben ser explícitamente justificadas.
A11 Impartial Reporting	A11 Reporte imparcial	El reporte debe evitar sesgos y distorsiones causadas por sentimientos personales para que refleje exactamente los hallazgos.
A12 Metaevaluation	A12 Metaevaluación	La evaluación en sí misma debe ser evaluada usando como base estándares pertinentes que permitan examinar sus fortalezas y debilidades.

Fuente: Elaboración propia a partir de *The Program Evaluation Standards*, Joint Committee on Standards for Educational Evaluations, 1994 (Traducción mía)

## B. Anexo B: Estándares de evaluación en el Sistema de las Naciones Unidas

Estándares de evaluación en el Sistema de las Naciones Unidas	
1. Marco institucional y gestión de la función de evaluación	
Marco institucional	
Estándar 1.1: Las organizaciones de las NU deben contar con un marco institucional adecuado para la gestión efectiva de su función de evaluación.	Promover una cultura que valore la evaluación como una base para el aprendizaje.



Estándares de evaluación en el Sistema de las Naciones Unidas	
	Facilitar un proceso de evaluación independiente e imparcial, garantizando que la función de evaluación se realice de manera independiente de otras funciones gerenciales. El jefe/a de la unidad de evaluación deberá reportar directamente al órgano de gobierno de la organización o al jefe/a de ésta.
Estándar 1.2: Las organizaciones de las NU deben elaborar una política de evaluación y actualizarla regularmente, teniendo en consideración las Normas y los Estándares de Evaluación del sistema de las NU.	La política de evaluación debe ser aprobada por los órganos de gobierno de las organizaciones y/o el jefe/a de éstas y debe ser coherente con las Normas de Evaluación de UNEG que sean aplicables, así como con los objetivos y estrategias corporativos. La política de evaluación debe incluir:
	una explicación clara del concepto y el rol de la evaluación al interior de la organización;
	una definición clara de los roles y responsabilidades de los profesionales directamente vinculados a la función de evaluación, de la alta gerencia y de los gerentes de programas;
	un énfasis en la necesidad de que los lineamientos de evaluación de la organización, sean adoptados;
	una explicación de la forma en que se priorizan y planifican las evaluaciones;
	una descripción de la forma en que se organizan, gestionan y presupuestan las evaluaciones;
	un énfasis en los requerimientos para el seguimiento de las evaluaciones; un enunciado claro en materia de divulgación y difusión.
Estándar 1.3: Las organizaciones de las NU deben asegurar que los planes de evaluación de actividades de evaluación sean presentados a sus órganos de gobierno y/o a los jefe/as de organizaciones para su revisión y/o aprobación.	4. Los órganos de gobierno y/o el jefe/a de la organización deben recibir no solamente el plan de evaluación sino también un informe de avance acerca de la ejecución tanto del plan de evaluación como de las recomendaciones que emanan de las evaluaciones.
Estándar 1.4: Las organizaciones de las NU deben garantizar la existencia de mecanismos apropiados para el seguimiento de las evaluaciones; y formular una política explícita en materia de divulgación.	5. Deben existir mecanismos apropiados para el seguimiento de las evaluaciones al interior de la organización, de tal suerte que garanticen que las recomendaciones de las evaluaciones sean debidamente utilizadas y oportunamente implementadas, y que sus hallazgos sean vinculados a actividades futuras.
	6. Una política de divulgación debe garantizar la difusión transparente de los resultados de las evaluaciones, lo cual incluye poner los informes a disposición general de los órganos de gobierno y del público, salvo en aquellos casos en que se requiera protección y confidencialidad razonables para algunas de las partes interesadas.
Gestión de la función de evaluación	
Estándar 1.5: El jefe/a de evaluación ejerce un rol de liderazgo en garantizar que la función de evaluación se encuentre plenamente operativa y que el trabajo en materia de evaluación, se lleve a cabo de acuerdo con los más altos estándares profesionales.	Dentro del marco institucional global, la gestión de la función de evaluación, encomendada al jefe/a de evaluación, debe garantizar que:
	una política de evaluación sea elaborada y actualizada regularmente;
	el presupuesto para evaluaciones sea administrado de manera eficiente;
	un plan de actividades, en materia de evaluación, sea elaborado una vez al año o cada dos años como parte del ciclo de planificación y presupuestación de la organización. Dicho plan priorizará aquellas áreas donde haya mayor requerimiento de una evaluación, y especificará los recursos adecuados para la planificación, realización y seguimiento de evaluaciones;

Estándares de evaluación en el Sistema de las Naciones Unidas	
	<p>metodologías de evaluación adecuadas sean adoptadas, desarrolladas y actualizadas en forma frecuente;</p> <p>las evaluaciones sean llevadas a cabo de acuerdo a estándares de calidad definidos y de una manera oportuna, para que sirvan como una herramienta valiosa para las partes interesadas y para los usuarios previstos;</p> <p>los informes a la alta dirección sean oportunos y relevantes a sus necesidades, impulsando con ello un proceso informado de toma de decisiones de gestión y de política institucional;</p> <p>informes periódicos de avance sobre la ejecución del plan de evaluación y/o la implementación de las recomendaciones que emanan de las evaluaciones ya realizadas, sean compilados y presentados a los órganos de gobierno y/o jefe/as de las organizaciones;</p> <p>las lecciones extraídas de las evaluaciones sean destiladas y difundidas según se juzgue conveniente.</p>
Estándar 1.6: El jefe/a de evaluación es responsable de asegurar la elaboración de lineamientos de evaluación.	
Estándar 1.7: El jefe/a de evaluación debe cerciorarse de que la función de evaluación sea dinámica, adaptándose a los nuevos desarrollos y a las necesidades cambiantes tanto dentro como fuera de la organización.	<p>En particular, la gestión de la función de evaluación debe incluir:</p> <p>generar conciencia y/o construir capacidad en evaluación;</p> <p>facilitar y gestionar redes de evaluación;</p> <p>diseñar e implementar metodologías y sistemas de evaluación;</p> <p>garantizar el mantenimiento de una memoria institucional de las evaluaciones, mediante mecanismos que sean amigables con el usuario;</p> <p>promover la compilación de lecciones de una manera sistemática.</p>
2. Competencias y ética	1. Todas las personas involucradas en el diseño, la realización y la gestión de actividades de evaluación deben aspirar a realizar un trabajo de alta calidad, que se rija por estándares profesionales y principios éticos y morales.
Estándar 2.1: Las personas involucradas en el diseño, la realización y la gestión de actividades de evaluación deben poseer competencias medulares en materia de evaluación.	
Estándar 2.2: Los evaluadores/as deben poseer educación, calificaciones y capacitación relevantes, en materia de evaluación.	
Estándar 2.3: Los evaluadores/as deben contar con experiencia de trabajo profesional relevante en materia de evaluación.	
Consideraciones éticas	
Estándar 2.5: Los evaluadores/as deben ser sensibles a las creencias, usos y costumbres y conducirse con integridad y honestidad en sus relaciones con todas las partes interesadas.	

Estándares de evaluación en el Sistema de las Naciones Unidas	
Estándar 2.6: Los evaluadores/as deben cerciorarse de que sus relaciones con las personas se caractericen por el respeto.	
Estándar 2.7: Los evaluadores/as deben proteger el anonimato y la confidencialidad de informantes particulares.	
Estándar 2.8: Los evaluadores/as son responsables de su desempeño y de su(s) producto(s).	
3. Realización de evaluaciones	
Estándar 3.1: La evaluación debe ser diseñada de manera tal que garantice información oportuna, válida, confiable y relevante sobre la intervención que está siendo evaluada.	
Estándar 3.2: Los términos de referencia deben precisar el propósito y describir el proceso y el producto de la evaluación.	
Estándar 3.3: El propósito y el contexto de la evaluación deben ser claramente descritos, proporcionando una justificación específica del por qué se emprenderá la evaluación en una fecha determinada.	
Estándar 3.4: La intervención que será evaluada debe ser claramente descrita.	
Estándar 3.5: Los objetivos específicos de la evaluación deben ser realistas y alcanzables, a la luz de la información que puede ser recolectada en el contexto de la intervención. El alcance de la evaluación debe ser también claramente definido.	
Estándar 3.6: El diseño de la evaluación debe detallar claramente los criterios a partir de los cuales se evaluará la intervención.	
Estándar 3.7: Las metodologías de evaluación deben ser lo suficientemente rigurosas como para evaluar la intervención y garantizar una evaluación exhaustiva, justa y libre de sesgos.	12. Los métodos de evaluación dependen de la información que se busque y del tipo de datos que se esté analizando. Los datos deben provenir de una variedad de fuentes a fin de garantizar la exactitud, validez y confiabilidad de los mismos, y que todas las personas afectadas y/o partes interesadas sean tomadas en cuenta. La metodología debe incidir explícitamente en los aspectos de género y de grupos subrepresentados.
	13. Deben reconocerse también, las limitaciones de los métodos de evaluación que fueron seleccionados.
Estándar 3.8: Una evaluación debe medir, en lo posible, el costo-efectividad.	

Estándares de evaluación en el Sistema de las Naciones Unidas	
Estándar 3.9: El diseño de la evaluación debe incluir, donde sea relevante, consideraciones sobre hasta qué punto el compromiso del sistema de las NU con el enfoque basado en los derechos humanos ha sido incorporado en el diseño de la intervención que será evaluada.	20. El diseño de la evaluación podría incluir además, algún proceso de revisión ética del diseño original de la intervención que será evaluada., que incluya lo siguiente: un balance de costos y beneficios para los participantes, incluyendo el potencial impacto negativo; las consideraciones éticas que subyacen a quién se incluye o excluye de la evaluación y cómo se lleva a cabo este proceso; el manejo de la privacidad y la confidencialidad; las prácticas para obtener acuerdos informados; la retroalimentación a los participantes; los mecanismos para determinar y monitorear la conducta y la práctica de los evaluadores/as y recolectores/as de datos.
Proceso	
Estándar 3.10: La relación entre el evaluador/a y la institución o instituciones que comisionan la evaluación tiene que caracterizarse, desde el principio, por el respeto mutuo y la confianza.	22. Los evaluadores/as deben consultar con la institución o instituciones que comisionan la evaluación acerca de decisiones contractuales tales como: confidencialidad, privacidad, comunicación y propiedad de los hallazgos e informes.
Estándar 3.11: Las partes interesadas deben ser consultadas en las etapas de planificación, diseño, realización y seguimiento de las evaluaciones.	
Estándar 3.12: Una revisión de pares o un grupo referencial integrado por expertos externos, puede ser especialmente valioso.	
Selección del equipo	
Estándar 3.13: Las evaluaciones deben ser llevadas a cabo por equipos de evaluación altamente calificados.	
Estándar 3.14: La composición de los equipos de evaluación debe ser balanceada en términos de género, geográficamente diversa e incluir profesionales de los países o regiones concernidos.	
Ejecución	
Estándar 3.15: Las evaluaciones deben llevarse a cabo de una manera profesional y ética.	
Informes	
Estándar 3.16: El informe final de evaluación debe contar con una estructura lógica, contener hallazgos basados en evidencia, conclusiones, lecciones y recomendaciones y estar libre de información irrelevante para el análisis general. El informe debe ser presentado en un formato que permita que la información sea comprensible y accesible.	37. El lector o lectora del informe de evaluación debe poder entender: el propósito de la evaluación; qué, exactamente, se evaluó; cómo se diseñó y llevó a cabo la evaluación; qué evidencias se encontró; qué conclusiones se extrajeron; qué recomendaciones se hicieron; qué lecciones se derivaron.

Estándares de evaluación en el Sistema de las Naciones Unidas	
	39. Los evaluadores/as deben permitir a todos los actores pertinentes, acceso a información procedente de la evaluación que sea apropiada, y, de ser posible, deben difundir activamente dicha información entre las partes interesadas. Las comunicaciones a una parte interesada deben incluir, siempre, los resultados importantes que pudieran afectar sus intereses. En todos los casos, los evaluadores/as deben esforzarse por presentar los resultados con la mayor claridad y sencillez posible, de manera tal que los clientes y otras partes interesadas puedan entender fácilmente el proceso y los resultados de la evaluación.
Seguimiento	
Estándar 3.17: La evaluación requiere de una respuesta explícita por parte de las autoridades de gobierno y gerencia afectadas por sus recomendaciones.	40. Al igual que en el caso de las Normas, dicha respuesta puede adoptar la forma de una respuesta de gerencia, un plan de acción y/o un acuerdo especificando claramente las responsabilidades y rendiciones de cuentas.
	41. Debe hacerse un seguimiento sistemático de la implementación de las recomendaciones de la evaluación que hubieran sido aceptadas por la gerencia y/o los órganos de gobierno.
	42. Debe realizarse, asimismo, un informe periódico acerca del estado de ejecución de las recomendaciones de la evaluación. Dicho informe debe ser presentado a los órganos de gobierno y/o el jefe/a de la organización.
4. Informes de evaluación	
Estándar 4.1: La portada y las páginas iniciales del informe deben ofrecer información básica clave.	
Estándar 4.2: El informe de evaluación debe contener un resumen ejecutivo.	2. Para facilitar la lectura, el resumen ejecutivo debe ser corto, de dos o tres páginas, y auto-contenido. El nivel de la información debe permitir al lector/a no letrado, una comprensión clara de los hallazgos, recomendaciones y lecciones de la evaluación.
	3. El resumen ejecutivo debe constar de: – una breve descripción de la intervención que está siendo evaluada; – el contexto, la situación actual y una descripción de la intervención en relación con otros asuntos vinculados; – el propósito de la evaluación; – los objetivos específicos de la evaluación; – el público objetivo del informe; – una breve descripción de la metodología, incluyendo la razón fundamental para haberla seleccionado, las fuentes de datos utilizadas, los métodos de recolección y análisis de datos empleados; y las principales restricciones relacionadas con la metodología; – los hallazgos y conclusiones más importantes; – las principales recomendaciones.
Estándar 4.3: La intervención que está siendo evaluada debe ser claramente descrita, incluyendo el modelo lógico y/o la cadena de resultados y el impacto esperados, la estrategia de ejecución y los supuestos centrales.	
Estándar 4.4: El rol y las contribuciones de las organizaciones de las NU y de otras partes interesadas de la intervención que está siendo evaluada deben describirse claramente.	

Estándares de evaluación en el Sistema de las Naciones Unidas	
Estándar 4.5: El propósito y el contexto de la evaluación deben consignarse en el informe.	
Estándar 4.6: El informe de evaluación debe ofrecer una explicación de los criterios de evaluación que fueron utilizados por los evaluadores/as	
Estándar 4.7: El informe de evaluación debe ofrecer una explicación clara de los objetivos específicos, así como del alcance de la evaluación.	
Estándar 4.8: El informe de evaluación debe indicar hasta qué punto fueron incorporadas consideraciones relevantes de género y de derechos humanos ahí donde fueran aplicables.	
Estándar 4.9: La metodología de evaluación aplicada debe describirse de una manera transparente, incluyendo cualesquier limitaciones inherentes a ella.	
Estándar 4.10: La evaluación debe proporcionar una descripción detallada de la participación de las partes interesadas.	17. Debe incluirse una descripción del nivel de participación de las partes interesadas en la evaluación, incluyendo la razón fundamental para haber seleccionado dicho nivel de participación en particular. Si bien no todas las evaluaciones pueden ser participativas en la misma medida, es importante otorgar la debida importancia a la participación de las partes interesadas, dado que dicha participación es crecientemente reconocida como un factor crucial en el uso de las conclusiones, recomendaciones y lecciones.
Estándar 4.11: El informe de evaluación debe incluir un análisis de la medida en que el diseño de la evaluación incluyó salvaguardas éticas donde fuera apropiado.	
Estándar 4.12: En la presentación de los hallazgos es necesario, en la medida de lo posible, medir los insumos, productos y efectos / impactos (o dar una explicación apropiada de por qué no se hace).	20. Los efectos e impactos deben incluir cualesquier resultados no intencionales, ya sean beneficiosos o perjudiciales. En adición, deberá incluirse cualesquier efectos multiplicadores o a nivel micro, de la intervención que está siendo evaluada. En la medida de lo posible, cada uno de éstos efectos deberá ser medido ya sea cuantitativa o cualitativamente. Al utilizar dichas mediciones, deberá hacerse referencia a parámetros referenciales.
Estándar 4.13: El análisis debe incluir una discusión apropiada de las contribuciones relativas, de las partes interesadas a los resultados.	
Estándar 4.14: Deben ser identificadas, en la medida de lo posible, las razones de los logros y las dificultades de la intervención que está siendo evaluada, especialmente los factores de restricción y habilitación.	26. El informe de evaluación debe ir más allá de una simple descripción de la implementación y de los efectos e incluir un análisis, basado en los hallazgos, de las causas subyacentes, las restricciones, las fortalezas que hay que capitalizar y las oportunidades. Deben identificarse y analizarse, en la medida de lo posible, los factores externos que contribuyen a los logros y dificultades, entre ellos la situación social, política o ambiental.

Estándares de evaluación en el Sistema de las Naciones Unidas	
Estándar 4.15: Las conclusiones tienen que ser sustentadas por hallazgos que sean consistentes con los datos recolectados y por la metodología; asimismo, deben aportar elementos de juicio conducentes a la identificación y/o solución de problemas o temas importantes.	
Estándar 4.16: Las recomendaciones deben estar firmemente basadas en evidencia y análisis y ser relevantes y realistas, con prioridades para la acción claramente enunciadas.	
Estándar 4.17: Las lecciones, al ser presentadas, deben incluir generalizaciones que trasciendan la intervención inmediata que está siendo evaluada poniendo énfasis en la relevancia más amplia que podrían tener.	
Estándar 4.18: Los anexos deben ser completos y relevantes.	

Fuente: Elaboración propia a partir de *Estándares de Evaluación en el Sistema de las Naciones Unidas*, Grupo de Evaluación de las Naciones Unidas (UNEG), 2005.





## C. Anexo C: Estándares de calidad para la evaluación del desarrollo del Comité de Ayuda al Desarrollo (CAD), de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), 2006

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006	
1. MOTIVO, PROPÓSITO Y OBJETIVOS DE UNA EVALUACIÓN	
1.1. El motivo de la evaluación.	Describe por qué y para quién se hace la evaluación, y por qué se realiza en un momento determinado.
1.2. El propósito de la evaluación.	El propósito de la evaluación está en línea con las funciones de aprendizaje y rendición de cuentas de las evaluaciones. Por ejemplo, los propósitos de una evaluación pueden ser:
	Contribuir a mejorar una política de ayuda, procedimiento o técnica.
	Considerar si es oportuno continuar o no un programa o proyecto.
	Rendir cuentas de los gastos en desarrollo a los actores implicados y a los contribuyentes.
1.3. Los objetivos de la evaluación.	Los objetivos de evaluación especifican lo que la evaluación pretende conseguir. Por ejemplo:
2. ALCANCE DE LA EVALUACIÓN	
2.1. Alcance de la evaluación.	El alcance de la evaluación se define claramente especificando las cuestiones que aborda, los fondos gastados hasta el momento, el período de tiempo, el tipo de intervención, la cobertura geográfica, los colectivos beneficiarios, así como cualquier otro elemento de la intervención de desarrollo que deba considerarse en la evaluación.

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006	
2.2. La lógica de la intervención y los resultados.	El informe de la evaluación describe brevemente y valora la lógica de la intervención, y distingue entre resultados a diferentes niveles: insumos, actividades, resultados e impactos. El informe también hace una valoración general de la lógica de la intervención.
2.3. Criterios de evaluación.	El informe de la evaluación aplica los cinco criterios del CAD para las evaluaciones de la cooperación al desarrollo: pertinencia, eficacia, eficiencia, impacto y sostenibilidad. Los criterios aplicados en una evaluación determinada se definen con claridad y sin ambigüedades. En caso de que no se haya aplicado un criterio o se hayan aplicado otros criterios adicionales se explica el porqué en el informe de evaluación.
2.4. Cuestiones de evaluación.	Las cuestiones preguntadas, incluyendo cualquier revisión de las preguntas originales, están documentadas en el informe de manera que los lectores sean capaces de valorar si el equipo de evaluación las ha respondido adecuadamente.
3. CONTEXTO	
3.1. El contexto de desarrollo y el contexto de la política.	El informe de evaluación hace una breve descripción del contexto político relevante para la intervención del desarrollo, de los documentos políticos, objetivos y estrategias de la agencia de desarrollo y de los socios.
	El contexto de desarrollo puede referirse a: la economía regional y nacional y a los niveles de desarrollo.
	El contexto de políticas puede referirse a: Estrategias de reducción de la pobreza, igualdad de género, protección del medio ambiente y derechos humanos.
3.2. El contexto institucional.	El informe de evaluación describe el contexto institucional y de los actores implicados relevantes para la intervención de desarrollo, de modo que su influencia pueda identificarse y valorarse.
3.3. El contexto sociopolítico.	El informe de evaluación describe el contexto sociopolítico en el que la intervención se desarrolla, así como su influencia en los resultados e impactos de la intervención de desarrollo.
3.4. Planes de ejecución.	El informe de evaluación describe los planes organizacionales establecidos para la ejecución de la intervención de desarrollo, incluidas las funciones de los donantes y socios.
4. METODOLOGÍA DE EVALUACIÓN	

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006	
4.1. Explicación de la metodología utilizada.	El informe de evaluación describe y explica la metodología y el proceso de evaluación, y analiza su validez y fiabilidad. Identifica todos los impedimentos encontrados, y sus implicaciones para la evaluación, incluyendo el impacto sobre su independencia. Detalla los métodos y técnicas usados para la recopilación y análisis de los datos y la información, justificando la alternativa elegida y explicando sus limitaciones y puntos débiles
4.2. Valoración de resultados.	Se especifican los métodos para valorar los resultados. Deben ser considerados los factores de atribución y de contribución/confusión. Si se han usado indicadores de base para la valoración de los resultados, estos deben ser específicos, medibles, alcanzables, relevantes y en tiempo limitado (SMART).
4.3. Consulta con los actores relevantes.	Los actores relevantes están implicados en el proceso de evaluación para identificar cuestiones y suministrar insumos para la evaluación. Han sido consultados donantes y socios. El informe de evaluación indica los actores consultados, los criterios con los que han sido seleccionados y describe su participación. Si no se ha consultado a todos los actores, se explican los métodos y razones seguidos para la selección de unos actores determinados.
4.4. Muestreo.	El informe de evaluación explica los criterios de selección de cualquier muestra e identifica las limitaciones relativas a su representatividad.
4.5. El equipo de evaluación.	El conjunto de los miembros del equipo de evaluación debe poseer una combinación de habilidades evaluadoras y conocimiento de los temas tratados, estar equilibrado en términos de género e incluir profesionales de los países o regiones implicados.
5. FUENTES DE INFORMACIÓN	
5.1. Transparencia de las fuentes de información.	El informe de evaluación describe con suficiente detalle las fuentes de información utilizadas (documentación, informantes clave, literatura, etc.), de modo que se puede valorar la adecuación de la información. Se incluyen listas completas de entrevistados y de documentos consultados, siempre que esto no entre en conflicto con el derecho a la privacidad y confidencialidad de los participantes.
5.2. Fiabilidad y exactitud de las fuentes de información.	Se comprueban y valoran críticamente las fuentes de información utilizadas y la validez de los datos, utilizando diversos métodos y fuentes de información.
6. INDEPENDENCIA	

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006	
6.1. Independencia del evaluador frente a los actores.	El informe de evaluación indica el grado de independencia de los evaluadores con respecto a las políticas, operaciones y las funciones de gestión de los que comisionan, implementan la intervención y de los beneficiarios. Los posibles conflictos de interés se abordan abierta y honradamente.
6.2. Proceso de evaluación, libre y abierto.	El equipo de evaluación puede trabajar libremente sin interferencias. Se asegura la cooperación y acceso a toda la información relevante. El informe de evaluación indica cualquier obstrucción que puede haber influido en el proceso de evaluación.
7. ÉTICA DE LA EVALUACIÓN	
7.1. Evaluación llevada a cabo de manera profesional y ética.	El proceso de evaluación tiene en cuenta el género, creencias, maneras y costumbres de todos los actores y se ha llevado a cabo con integridad y honradez. Se protegen los derechos y el bienestar de los participantes en la evaluación. El anonimato y la confidencialidad de los informantes individuales deberán estar protegidos cuando son solicitados o establecidos por ley.
7.2. Reconocimiento de los desacuerdos entre el equipo de evaluación.	Los miembros del equipo de evaluación deben tener la oportunidad de desvincularse de cualquier juicio y recomendación. Cualquier diferencia de opinión no resuelta entre los miembros del equipo deberá constar en el informe.
8. CALIDAD	
8.1. Incorporación de los comentarios de los actores involucrados.	Los actores han tenido la oportunidad de comentar los resultados, conclusiones, recomendaciones y lecciones aprendidas. El informe de evaluación refleja esos comentarios y reconoce cualquier desacuerdo sustantivo. En las discrepancias sobre hechos que pueden ser verificados, los evaluadores deben investigar y cambiar el borrador cuando sea necesario. En el caso de opiniones o interpretaciones, los comentarios de los actores deben ser reproducidos textualmente, como anexo, siempre que esto no entre en conflicto con los derechos y bienestar de los participantes.
8.2. Control de calidad.	Se utiliza el control de calidad durante todo el proceso de evaluación. En función del alcance de la evaluación y de su complejidad, el control de calidad se puede realizar de manera interna o externa, a través de un examen de pares o mediante un grupo de referencia. Los controles de calidad cumplen el principio de independencia del evaluador.
9. RELEVANCIA DE LOS RESULTADOS DE LA EVALUACIÓN	

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006	
9.1. Formulación de los resultados de la evaluación.	Los resultados de la evaluación son relevantes para el objeto evaluado y el propósito de la evaluación. Los resultados deben proceder claramente de las cuestiones de evaluación y del análisis de datos, mostrando una clara línea de evidencia que apoye las conclusiones. Se explica cualquier discrepancia entre lo planeado y la implementación real del objeto evaluado.
9.2. Evaluación ajustada al tiempo y al presupuesto.	La evaluación se realiza y los resultados se presentan en un tiempo y manera acorde con el propósito de la evaluación. En el informe se explica cualquier modificación imprevista del cronograma o del presupuesto. Se explica cualquier discrepancia entre la evaluación planeada y la evaluación realizada.
9.3. Recomendaciones y lecciones aprendidas.	Las recomendaciones y lecciones aprendidas son relevantes, se dirigen a los usuarios adecuados y son susceptibles de llevarse a cabo bajo la responsabilidad de los usuarios. Las recomendaciones son propuestas opinables y las lecciones aprendidas son generalizaciones de las conclusiones, aplicables a un ámbito más amplio.
9.4. Uso de la evaluación.	La evaluación requiere un reconocimiento explícito y una respuesta de los gestores sobre la intención de poner en marcha y dar seguimiento a los resultados de la evaluación. La unidad gestora de la evaluación debe asegurar la difusión sistemática, la conservación y la gestión de los resultados de las evaluaciones para asegurar su accesibilidad y maximizar los beneficios de los hallazgos de la evaluación.
10. EXHAUSTIVIDAD	
10.1. Cuestiones de evaluación a las que responden las conclusiones.	En el informe de evaluación responde a todas las cuestiones y necesidades de información detalladas en el alcance de la evaluación. Cuando no es posible se explican las razones.
10.2. Claridad del análisis.	El análisis está estructurado de forma lógica. Los datos y la información se presentan, analizan e interpretan sistemáticamente. Los resultados y conclusiones están identificados claramente y son la conclusión lógica del análisis de los datos y de la información. Las presunciones e hipótesis subyacentes se han explicitado y se han tenido en cuenta.
10.3. Distinción entre conclusiones, recomendaciones y lecciones aprendidas.	En los informes de evaluación se debe distinguir claramente entre resultados, conclusiones, recomendaciones. La evaluación los presenta de manera separada y con una distinción entre ellas clara y lógica. Las conclusiones están basadas en resultados y análisis. Las recomendaciones y lecciones aprendidas provienen de manera lógica de las conclusiones.

Organización para la Cooperación y el Desarrollo Económicos (OCDE), Comité de Ayuda al Desarrollo (CAD) 2006		
10.4.	Claridad y representatividad del resumen ejecutivo.	El informe de evaluación se acompaña de un resumen ejecutivo, en el que se suministra una visión global del informe y se subrayan las principales conclusiones, recomendaciones y lecciones aprendidas.

Fuente: Elaboración propia con base en: Organización para la Cooperación y el Desarrollo Económicos –OCDE-, Comité de Ayuda al Desarrollo –CAD-, Estándares de Calidad para la Evaluación del Desarrollo, 2006.

## Anexo D: Tablas de calificación para cada una de las evaluaciones

Indicadores de goce efectivo de derechos para población desplazada						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1. MARCO INSTITUCIONAL	23,33		7	20%	1,4	
1.1 El proceso de evaluación ha sido independiente e imparcial	70	30%	21			Si bien las firmas contratadas no tienen ninguna relación con el programa, las firmas Econometría y SEI participan en 3 de las 5 evaluaciones analizadas, por lo que no se garantiza diversidad en los posibles evaluadores. Se debe reconocer, sin embargo, que el mercado de evaluadores en Colombia es pequeño y que además el respaldo financiero que se exige para participar en la licitación es muy alto para que firmas pequeñas tengan la posibilidad.
1.2 Enfoque asociativo		30%	0			La evaluación se aplica de arriba hacia abajo y no evidencia la participación de distintos actores, aparte del gobierno, la firma contratada para realizar la evaluación y los beneficiarios tienen un rol

Indicadores de goce efectivo de derechos para población desplazada						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						pasivo (dan información) pero no participan del proceso de evaluación
1.3 Conflicto de intereses		40%	0			No se evidencia ningún conflicto de interés ni en el diseño ni en la ejecución de la evaluación. Es posible que esta información esté consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.
2. ENFOQUE Y METODOLOGÍA	70		35	25%	8,75	
2.1 Selección del enfoque y de la metodología	60	50%	30			La metodología está claramente explicada y coincide con el enfoque definido. Los procesos de recolección de información están detallados



Indicadores de goce efectivo de derechos para población desplazada						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
2.2 Limitaciones de los métodos de evaluación	80	50%	40			Este parámetro se ha evaluado positivamente porque en la evaluación se reconocen algunas de las imitaciones de la metodología: por ejemplo que algunos indicadores cambiaron en la forma de medición respecto a la línea base, por lo que no son comparables con los datos de 2008, y se consideran por lo tanto una nueva línea base. Es el caso de indicadores como generación de ingresos. Adicionalmente, se señala que la concentración por razones de eficiencia de la selección de la muestra de hogares en las cabeceras municipales del país y áreas periféricas, impide conocer las situación de los desplazados que han ubicado en zonas rurales o que han retornado a su sitio de residencia en dicha zona, agravada por el hecho de que no se puede estimar que proporción de dicha población ha elegido este tipo de lugar de residencia

Indicadores de goce efectivo de derechos para población desplazada						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3. INFORME FINAL	63,6		12,84	30%	3,852	
3.1 Conclusiones justificadas	65	25%	16,25			Las conclusiones se justifican con base en los datos y análisis contenidos en el informe, pero no se puede perder de vista el problema de comparabilidad que existe entre los datos
3.2 Imparcialidad	65	25%	16,25			A pesar de que se reconoce la falta de comparabilidad entre la línea base de 2008 y los datos levantados en 2010, se recomienda aceptar esta dificultad metodológica y se concluye que la mayoría de los indicadores han mejorado
3.3 Claridad	59	20%	11,8			Aunque el informe está escrito de forma estructurada tiene algunos apartes referidos a metodología y a cálculos que resultan muy especializados para ser entendibles para todo público
3.4 Impactos no intencionales	59	10%	5,9			No se registraron impactos no intencionales, pero esto puede obedecer a la metodología empleada porque limita la posibilidad de hacer otros análisis

Indicadores de goce efectivo de derechos para población desplazada						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						"efecto túnel"
3.5 Factores externos y de contexto	70	20%	14			La documentación del contexto es detallada y aporta elementos importantes para comprender la complejidad del fenómeno en el marco de la evaluación
4. IMPACTO DE LA EVALUACIÓN	16,67		5,83333	25%	1,45833	
4.1 Difusión transparente y sistemática de los resultados de la evaluación	0	25%	0			No se evidencia ninguna acción de socialización o difusión de los resultados del informe con actores ajenos al equipo evaluador.
4.2 Uso de la evaluación	50	35%	17,5			Tampoco se evidencia que la evaluación se haya usado para la toma de decisiones, la rendición de cuentas o el control ciudadano
4.3 Metaevaluación	0	40%	0			No se realizó ninguna metaevaluación de esta evaluación
PUNTAJE GENERAL				100%	3,86508	

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1. MARCO INSTITUCIONAL	46,67		14	20%	2,8	
1.1 El proceso de evaluación ha sido independiente e imparcial	70	30%	21			Esta es la única firma de toda la muestra que no es un consultor de evaluación común o de los más contratados en el mercado cerrado de consultores del país, lo cual se considera positivo pero no es una garantía de independencia por tal razón se pondera moderadamente.
1.2 Enfoque asociativo	70	30%	21			En esta evaluación se tuvieron en cuenta actores distintos a los beneficiarios que tienen relación con el programa como Constructores, Cajas de

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						Compensación Familiar y Entes Territoriales, a través de entrevistas semiestructuradas y grupos focales
1.3 Conflicto de intereses	0	40%	0			No se evidencia ningún conflicto de interés ni en el diseño ni en la ejecución de la evaluación. Es posible que esta información esté consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.
2. ENFOQUE Y METODOLOGÍA	42,5		21,25	25%	5,3125	

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
2.1 Selección del enfoque y de la metodología	85	50%	42,5			La metodología está claramente explicada y coincide con el enfoque definido. Los procesos de recolección de información están detallados. En este caso, aunque también se utilizó la metodología cuasi-experimental, ésta se complementó con metodologías cualitativas (grupos focales), que fueron las que arrojaron los retrasos o impactos negativos o no esperados del programa
2.2 Limitaciones de los métodos de evaluación	0	50%	0			En la evaluación no se identifican las imitaciones de la metodología, sin embargo sí se señalan las dificultades enfrentadas durante el

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						proceso de recolección de información
3. INFORME FINAL	56		11	30%	3,3	
3.1 Conclusiones justificadas	80	25%	20			Las conclusiones se justifican con base en los datos y análisis contenidos en el informe y realizan aportes relevantes para la gestión del programa, sobre todo en lo que se refiere a los resultados encontrados tras la aplicación de metodologías cualitativas que indagaron sobre la percepción de los beneficiarios. Adicionalmente, destaca los aspectos no favorables de la ejecución del programa

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3.2 Imparcialidad	60	25%	15			No se evidencian sesgos en la evaluación que afecten la presentación de resultados
3.3 Claridad	60	20%	12			Aunque el informe está escrito de forma estructurada tiene algunos apartes referidos a metodología y a cálculos que resultan muy especializados para ser entendibles para todo público
3.4 Impactos no intencionales	80	10%	8			Aunque no se registran como impactos no intencionales, gracias a la aplicación de metodologías cualitativas, se pudo establecer que la consecución de la vivienda podría tener un efecto negativo en los procesos de organización y



Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						participación comunitaria, pues al lograr el objetivo que unía a los beneficiarios los lazos se debilitaron
3.5 Factores externos y de contexto	0	20%	0			El contexto de la intervención no se desarrolla y no se tienen en cuenta factores externos
4. IMPACTO DE LA EVALUACIÓN	20		7	25%	1,75	
4.1 Difusión transparente y sistemática de los resultados de la evaluación	0	25%	0			No se evidencia ninguna acción de socialización o difusión de los resultados del informe con actores ajenos al equipo evaluador.

Evaluación final y de impacto sobre la ejecución del Programa de Vivienda de Interés Social Urbana -VISU (1483)						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
4.2 Uso de la evaluación	60	35%	21			Tampoco se evidencia que la evaluación se haya usado para la toma de decisiones, la rendición de cuentas o el control ciudadano. Sin embargo, las recomendaciones se dividen por tipo de actor y por etapa del programa, lo cual podría contribuir al uso de la evaluación por parte de distintos actores.
4.3 Metaevaluación	0	40%	0			No se realizó ninguna metaevaluación de esta evaluación
PUNTAJE GENERAL				100%	3,29063	

Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1. MARCO INSTITUCIONAL	23,33		7	20%	1,4	
1.1 El proceso de evaluación ha sido independiente e imparcial	70	30%	21			La evaluación fue contratada con el Centro Nacional de Consultoría, que si bien solo es el evaluador de este programa en la muestra definida, ha participado en al menos 8 evaluaciones de las contratadas por el DNP. Se debe reconocer, sin embargo, que el mercado de evaluadores en Colombia es pequeño y que además el respaldo financiero que se exige para participar en la licitación es muy alto para que firmas pequeñas tengan la posibilidad.
1.2 Enfoque asociativo	0	30%	0			La evaluación se aplica de arriba hacia abajo y no evidencia la participación de distintos actores, aparte del gobierno, la firma contratada para realizar la evaluación y los beneficiarios tienen un rol pasivo (dan información) pero no participan del proceso de evaluación

Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1.3 Conflicto de intereses	0	40%	0			No se evidencia ningún conflicto de interés ni en el diseño ni en la ejecución de la evaluación. Es posible que esta información esté consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.
2. ENFOQUE Y METODOLOGÍA	0		0	25%	0	
2.1 Selección del enfoque y de la metodología		50%	0			La metodología está claramente explicada y coincide con el enfoque definido. En este caso, aunque también se utilizó la metodología cuasi-experimental (medición estructurada), que se complementó con metodologías cualitativas (medición no estructurada), para captar información complementaria y de profundización, mediante diálogos y discusiones colectivas e individuales, con diversos tipos de actores del programa.
2.2 Limitaciones de los métodos de evaluación	0	50%	0			En la evaluación no se identifican las imitaciones de la metodología.
3. INFORME FINAL	37		8,65	30%	2,595	

Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3.1 Conclusiones justificadas	65	25%	16,25			Las conclusiones se justifican con base en los datos y análisis contenidos en el informe, pero no se puede perder de vista el problema de comparabilidad que existe entre los datos
3.2 Imparcialidad	60	25%	15			No se evidencian sesgos en la evaluación que afecten la presentación de resultados
3.3 Claridad	60	20%	12			Aunque el informe está escrito de forma estructurada tiene algunos apartes referidos a metodología y a cálculos que resultan muy especializados para ser entendibles para todo público
3.4 Impactos no intencionales	0	10%	0			No se registraron impactos no intencionales, pero esto puede obedecer a la metodología empleada porque limita la posibilidad de hacer otros análisis "efecto túnel"

Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3.5 Factores externos y de contexto	0	20%	0			En la evaluación no se identifican ni se analizan, los factores externos que en este caso afectaron la validez de la evaluación debido a una coyuntura política particular. Resulta preocupante que en el informe final de evaluación no se haga referencia a las razones y se trate de remediar con un tecnicismo estadístico la situación.
4. IMPACTO DE LA EVALUACIÓN	23,33		8,16667	25%	2,04167	
4.1 Difusión transparente y sistemática de los resultados de la evaluación	0	25%	0			No se evidencia ninguna acción de socialización o difusión de los resultados del informe con actores ajenos al equipo evaluador.
4.2 Uso de la evaluación	70	35%	24,5			Tampoco se evidencia que la evaluación se haya usado para la toma de decisiones, la rendición de cuentas o el control ciudadano. Sin embargo, con base en las recomendaciones y en lo que luego ocurrió con el programa se infiere que éstas fueron tenidas en cuenta para el ajuste del programa
4.3 Metaevaluación	0	40%	0			No se realizó ninguna metaevaluación de esta evaluación

Evaluación del Programa Familias en Acción - Población en Situación de Desplazamiento						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
PUNTAJE GENERAL				100%	1,50917	

Programa Familias en Acción: Impactos en capital humano - Evaluación rural						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1. MARCO INSTITUCIONAL	23,33		7	20%	1,4	
1.1 El proceso de evaluación ha sido independiente e imparcial	70	30%	21			La evaluación fue contratada con la misma Unión Temporal conformada por el Institute of Fiscal Studies, Econometría y Sistemas Especializados de Información S.A. que realiza otra de las evaluaciones de la muestra y cuyas firmas participan en 3 de las 5 evaluaciones analizadas. Se debe reconocer, sin embargo, que el mercado de evaluadores en Colombia es pequeño y que además el respaldo financiero que se exige para participar en la licitación es muy alto para que firmas pequeñas tengan la posibilidad.

Programa Familias en Acción: Impactos en capital humano - Evaluación rural						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1.2 Enfoque asociativo	0	30%	0			La evaluación se aplica de arriba hacia abajo y no evidencia la participación de distintos actores, aparte del gobierno, la firma contratada para realizar la evaluación y los beneficiarios tienen un rol pasivo (dan información) pero no participan del proceso de evaluación
1.3 Conflicto de intereses	0	40%	0			No se evidencia ningún conflicto de interés ni en el diseño ni en la ejecución de la evaluación. Es posible que esta información esté consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.
2. ENFOQUE Y METODOLOGÍA	35		17,5	25%	4,375	
2.1 Selección del enfoque y de la metodología	70	50%	35			La metodología está claramente explicada y coincide con el enfoque definido. En este caso, aunque también se utilizó la metodología cuasi-experimental (medición estructurada), que se complementó con metodologías cualitativas (medición no estructurada), para captar información complementaria y de profundización, mediante diálogos y discusiones



Programa Familias en Acción: Impactos en capital humano - Evaluación rural						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						colectivas e individuales, con diversos tipos de actores del programa.
2.2 Limitaciones de los métodos de evaluación	0	50%	0			En la evaluación no se identifican las imitaciones de la metodología
3. INFORME FINAL	37		8,65	30%	2,595	
3.1 Conclusiones justificadas	65	25%	16,25			Las conclusiones se justifican con base en los datos y análisis contenidos en el informe, pero no se puede perder de vista el problema de comparabilidad que existe entre los datos
3.2 Imparcialidad	60	25%	15			No se evidencian sesgos en la evaluación que afecten la presentación de resultados
3.3 Claridad	60	20%	12			Aunque el informe está escrito de forma estructurada tiene algunos apartes referidos a metodología y a cálculos que resultan muy especializados para ser entendibles para todo

Programa Familias en Acción: Impactos en capital humano - Evaluación rural						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						público
3.4 Impactos no intencionales	0	10%	0			No se registraron impactos no intencionales, pero esto puede obedecer a la metodología empleada porque limita la posibilidad de hacer otros análisis "efecto túnel"
3.5 Factores externos y de contexto	0	20%	0			En la evaluación no se identifican ni se analizan, los factores externos que en este caso afectaron la validez de la evaluación debido a una coyuntura política particular. Resulta preocupante que en el informe final de evaluación no se haga referencia a las razones y se trate de remediar con un tecnicismo estadístico la situación.
4. IMPACTO DE LA EVALUACIÓN	23,33		8,16667	25%	2,04167	
4.1 Difusión transparente y sistemática de los resultados de la evaluación	0	25%	0			No se evidencia ninguna acción de socialización o difusión de los resultados del informe con actores ajenos al equipo evaluador.

Programa Familias en Acción: Impactos en capital humano - Evaluación rural						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
4.2 Uso de la evaluación	70	35%	24,5			Tampoco se evidencia que la evaluación se haya usado para la toma de decisiones, la rendición de cuentas o el control ciudadano. Sin embargo, con base en las recomendaciones y en lo que luego ocurrió con el programa se infiere que éstas fueron tenidas en cuenta para el ajuste del programa
4.3 Metaevaluación	0	40%	0			No se realizó ninguna metaevaluación de esta evaluación
PUNTAJE GENERAL				100%	2,60292	

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1. MARCO INSTITUCIONAL	23,33		7	20%	1,4	

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1.1 El proceso de evaluación ha sido independiente e imparcial	70	30%	21			La evaluación fue contratada con la misma Unión Temporal conformada por el Institute of Fiscal Studies, Econometría y Sistemas Especializados de Información S.A. que realiza otra de las evaluaciones de la muestra y cuyas firmas participan en 3 de las 5 evaluaciones analizadas. Se debe reconocer, sin embargo, que el mercado de evaluadores en Colombia es pequeño y que además el respaldo financiero que se exige para participar en la licitación es muy alto para que firmas pequeñas tengan la posibilidad.
1.2 Enfoque asociativo	0	30%	0			La evaluación se aplica de arriba hacia abajo y no evidencia la participación de distintos actores, aparte del gobierno, la firma contratada para realizar la evaluación y los beneficiarios tienen un rol pasivo (dan información) pero no participan del proceso de evaluación

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
1.3 Conflicto de intereses	0	40%	0			No se evidencia ningún conflicto de interés ni en el diseño ni en la ejecución de la evaluación. Es posible que esta información esté consignada en las actas de las reuniones del equipo evaluador a las que no fue posible acceder.
2. ENFOQUE Y METODOLOGÍA	35		17,5	25%	4,375	

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
2.1 Selección del enfoque y de la metodología	30	50%	15			La metodología está claramente explicada. Los procesos de recolección de información están detallados. En este caso, se tuvo que cambiar el diseño experimental por uno cuasi-experimental, debido a que el grupo de control no había sido elegido aleatoriamente, para medir los impactos atribuibles al Programa. Adicionalmente, se indagó por dos tipos de impactos: en los beneficiarios directos y los beneficios indirectos del programa sobre las personas residentes en la zona de influencia de los proyectos realizados en barrios de bajos ingresos. No obstante, la metodología no coincide con el enfoque definido para la evaluación en tanto no consideró los problemas en la operación del programa en los distintos territorios por lo que la metodología tuvo que ser ajustada, comprometiendo la validez de los resultados. Es por ello que se ha ponderado negativamente.

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
2.2 Limitaciones de los métodos de evaluación	40	50%	20			En la evaluación no se identifican las limitaciones de la metodología, sin embargo sí se señalan las dificultades enfrentadas durante el proceso de recolección de información que consistieron básicamente en: la ausencia de aleatoriedad de la selección de beneficiarios (ninguna de las muestras analizadas el proceso de selección de beneficiarios correspondió a un proceso aleatorio), las diferencias en la Línea de Base entre la población de personas seleccionadas inicialmente como beneficiarios y los participantes que finalmente trabajaron en las obras, después de las demoras en su iniciación, el tamaño del grupo de control terminó siendo la mitad del beneficiario, se tuvo que excluir un municipio, San Vicente del Caguán, como un caso extremo porque se alejó de la forma en la cual debía operar el Programa. Adicionalmente, 47 de los 155 proyectos que se encuestaron en línea de base fueron eliminados porque no se realizaron. Por

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
						estas razones el criterio ha sido ponderado negativamente.
3. INFORME FINAL	30		6,6	30%	1,98	



Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3.1 Conclusiones justificadas	0	25%	0			Las conclusiones se justifican con base en los datos y análisis contenidos en el informe. Sin embargo debido a los ajustes metodológicos los datos no parecen confiables y por lo tanto las conclusiones pueden ser cuestionadas. Es por ello que el sub-criterio ha recibido una baja calificación
3.2 Imparcialidad	60	25%	15			No se evidencian sesgos en la evaluación que afecten la presentación de resultados
3.3 Claridad	40	20%	8			Aunque el informe está escrito de forma estructurada tiene algunos apartes referidos a metodología y a cálculos que resultan muy especializados para ser entendibles para todo público. Adicionalmente, los cambios en la metodología resultan confusos.
3.4 Impactos no intencionales	0	10%	0			No se registraron impactos no intencionales, pero esto puede obedecer a la metodología empleada porque limita la posibilidad de hacer otros análisis "efecto túnel"

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
3.5 Factores externos y de contexto	50	20%	10			En la evaluación se identifican los factores externos que en este caso afectaron la recolección de datos de la evaluación debido a una demora en el inicio del programa. Sin embargo, no se analizan los factores externos sino que se tratan de resolver con un ajuste metodológico (estadístico)
4. IMPACTO DE LA EVALUACIÓN	0		0	25%	0	No se registraron impactos no intencionales, pero esto puede obedecer a la metodología empleada porque limita la posibilidad de hacer otros análisis "efecto túnel"
4.1 Difusión transparente y sistemática de los resultados de la evaluación	0	25%	0			No se evidencia ninguna acción de socialización o difusión de los resultados del informe con actores ajenos al equipo evaluador.
4.2 Uso de la evaluación	0	35%	0			Tampoco se evidencia que la evaluación se haya usado para la toma de decisiones, la rendición de cuentas o el control ciudadano. Pero llama la atención que esta evaluación se publicó en el año 2006 y el programa se había suspendido en 2003. Por esta razón se pondera negativamente.

Programa Empleo en Acción: Impactos sobre los jóvenes y sus familias						Observaciones
DETALLE	POR SUB-CRITERIO			POR CRITERIO		
	Nota	Peso	Puntaje	Peso	Puntaje	
4.3 Metaevaluación	0	40%	0			No se realizó ninguna metaevaluación de esta evaluación
PUNTAJE GENERAL				100%	1,93875	

## Bibliografía

Aguilar, L. F. (2000). La hechura de las políticas. Segunda Antología. México: Miguel Ángel Porrúa.

Aguilar, L. F. (2000). La implementación de las políticas. Cuarta Antología. México: Miguel Ángel Porrúa.

Amar, A. & Berthier, L. (s.f.) Le Nouveau Management Public: Avantages et Limites”. Revue du réseau RECEMAP. N° 13. Consultada en línea el 30/10/2010.

Atrio, J. L. & Piccone, M. S. (2008). De la Administración Pública a la Gerencia Pública. El porqué de la necesidad de gestionar la transición. Reforma y Democracia. Caracas: CLAD. No. 42.

Banco Mundial. (2007). Diagnóstico del Sistema Nacional de SyE de Colombia: SINERGIA. Grupo Independiente de Evaluación. Serie de documentos de trabajo. N° 17. Washington, D.C: Mackay, Keith (Autor principal).

Boussaguet, L., Jacquot, S. y Ravinet, P. (Directores) González, A., Jolly, J., Herrán, V. & Soto, D. (Traductores), (2009). Diccionario de políticas públicas. Bogotá: Universidad Externado de Colombia.

Bresser-Pereira, L. C. (2001, septiembre). Reforma de la nueva gestión pública: ahora en la agenda de América Latina, sin embargo.... En *International Journal of Political Studies*, No. 3 (pp. 143- 166). Conferencia organizada por Hellmutt Wolman, Evaluando la nueva gestión pública en el mundo. Río de Janeiro, Brasil.

Briceño, B., Cuesta, L. Attanasio, O. (2011). "Behind the scenes: experience managing and conducting large impact evaluations in Colombia." *Journal of Development Effectiveness*, 3(4), 470-501

Centro Latinoamericano de Administración para el Desarrollo CLAD. (1998). Una Nueva Gestión Pública para América Latina. Documento preparado por el Consejo Científico del CLAD y aprobado en su sesión del día 14 de octubre de 1998 por el Consejo Directivo del CLAD.

Constitución Política de Colombia, 1991.

Cooksy, L. and Caracelli, V. (2005) "Quality, Context, and Use: Issues in Achieving the Goals of Metaevaluation", *American Journal of Evaluation* 26: 31-42.

Cuervo, J. I. (2007). Ensayos sobre Políticas Públicas. Bogotá: Universidad Externado de Colombia.

Cunill, N., & Ospina S. (2008). Fortalecimiento de los sistemas de monitoreo y evaluación (M&E) en América Latina: Informe comparativo de 12 países. Washington, DC, Banco Mundial, CLAD.

Departamento Nacional de Planeación (1994). La evaluación de resultados en el sector público a nivel nacional (Documento CONPES 2688). Bogotá: DECER-DNP.

\_\_\_\_\_. (1995). Gestión pública orientada a resultados (Documento CONPES 2790). Bogotá: DEE-DNP.

\_\_\_\_\_. (1996). Hacia una gestión pública orientada a resultados. Bogotá: Editorial Lito Camargo.

\_\_\_\_\_. (2003, octubre). Renovación de la Administración Pública (Documento CONPES 3248). Bogotá: DNP.

\_\_\_\_\_. (2004, junio). Renovación de la Administración Pública. Gestión por Resultados y Reforma del Sistema Nacional de Evaluación (Documento CONPES 3294). Bogotá: DNP.

\_\_\_\_\_. (2006). “Programa Familias en Acción. Resultados de impacto un año después de implementado el programa” en Evaluación de Políticas Públicas, No. 4. Dirección de Evaluación de Políticas Públicas, Departamento Nacional de Planeación, Bogotá D.C.

DNP. (2008). “Programa Familias en Acción. Impactos en capital humano y evaluación beneficio-costos del programa” en Evaluación de Políticas Públicas, No. 4. Dirección de Evaluación de Políticas Públicas, Departamento Nacional de Planeación, Bogotá D.C.

\_\_\_\_\_. (2009). Guía metodológica para la formulación de indicadores. Bogotá: DNP.

\_\_\_\_\_. (2010, abril). 15 años del Sistema Nacional de Evaluación y Gestión de Resultados –Sinergia–. Una mirada desde las evaluaciones de política pública más relevantes. Bogotá: DNP.

\_\_\_\_\_. (2012). Guía Metodológica para el Seguimiento al Plan Nacional de Desarrollo y la Evaluación de Políticas Estratégicas. Serie de Guías Metodológicas Sinergia. Dirección de Evaluación de Políticas Públicas, Departamento Nacional de Planeación, Bogotá D.C.

Guba, Egon & Lincoln, Yvonna (1989). The Coming of Age Evaluation. En: E. Guba, y Y. Lincoln, Fourth generation Evaluation (pp. 21 – 49). Newbury Park. California.

Hassenteufel, P. (2005). De la comparaison internationale à la comparaison transnationale. Les déplacements de la construction d'objets comparatifs en matière de politiques publiques. *Revue Française de Science Politique*, 55 (1), 113 – 132.

Henry, G. T. and Mark, M. M. (2003). "Beyond use: Understanding evaluation's influence on attitudes and actions", *American Journal of Evaluation* 24: 293-314.

House, E.R., (1993) "Professional Evaluation", Newbury Park: Sage.

King, G., Keohane, R., Verba, S. (1994) "Designing Social Inquiry: Scientific Inference in Qualitative Research" Princeton, Princeton University Press.

Kothari, C. (1990) "Research Methodology: Methods and Techniques" New Delhi, New Age International.

Maxwell, J. (2005) "Qualitative Research Design: An Interpretative Approach", 2nd Edition, London, Sage Publications Inc.

Hintze, S. (2001, noviembre) Reflexiones sobre el conflicto y la participación en la evaluación de políticas sociales. Texto basado en la ponencia presentada en el VI Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública. Buenos Aires.

Larbi, G. (1999) The new public management approach and crisis states. Geneva: United Nations Research Institute for Social Development.

Lozano, Alejandro. (2008a) “Aspectos sobre política pública. Ejercicios e instrumentos para el análisis”

\_\_\_\_\_. (2008b). Dispositivos de publicación y referenciales de política pública. Caso: Parque Tercer Milenio. Universidad Nacional de Colombia. Bogotá. Febrero de 2008.

Majone, Giandomenico (1992). Los usos del análisis de políticas. En: L. Aguilar, La hechura de las políticas (pp. 341 – 366). México: Porrúa.

Meny, Y., & Thoenig J. C. (1992). Las políticas públicas. Barcelona: Ariel. Ciencia Política.

Muller, Pierre (2008) “Las Políticas Públicas”. Jean-François Jolly, Carlos Salazar (Traducción) 2da Edición. Estudios de caso en Políticas Públicas. Universidad Externado de Colombia. Bogotá.

Organización de Cooperación y Desarrollo Económicos, OCDE (Comité de Ayuda al Desarrollo, DAC) (2010) “Estándares de Calidad para la Evaluación del Desarrollo”, “Principles for the Evaluation of Development Assistance”, in OECD (DAC) Principles for



Effective Aid, OECD (DAC), Paris. Publicado originalmente por la OCDE en inglés y francés bajo los títulos, respectivamente: DAC Guidelines and Reference Series. Quality Standards for Development Evaluation, Lignes directrices et ouvrages de référence du CAD. Normes de qualité pour l'évaluation du développement.

Palumbo, D., Fawcett, S., y Wright, P. (Editor). (1981). Evaluating and optimizing Public Policy. Massachusetts: University of Kansas. Lexington Books.

Pierson, Paul (2004). Introduction. Chapter One. En P. Pierson, History, institutions, and social analysis Positive feedback and path dependence (pp. 1 - 53). Princeton University Press.

Pressman, J. & Wildavsky, A (1973). Implementation. Berkeley.

Rojas, F., Mackay, K., Matsuda, Y., Shepherd, G., Del Villar, A., Zaltsman, A., y Krause, P. (2005). "Chile: Análisis del Programa de Evaluación del Gasto Público". Estudios del Banco Mundial sobre América Latina y el Caribe. Banco Mundial, Unidad de Reducción de la Pobreza y Gestión Económica. Washington, D.C. Diciembre 2005.

Rossi, P.H., Freeman, H.E., (1989), Evaluation: A Systematic Approach, Beverly Hills: Sage.

Rossi, P., Lipsey, M. and Freeman, H. (2004) Evaluation: a systematic approach, 7th Edition, California, Sage.

Roth, André-Noël (2002). Políticas Públicas: Formulación, implementación y evaluación. Bogotá: Ediciones Aurora.

Roth, André-Noël (2007). Enfoques y teorías para el análisis de las políticas públicas, cambio de la acción pública y transformaciones del Estado. En Cuervo, J., Ensayos sobre Políticas Públicas (pp. 29-63). Bogotá: Universidad Externado de Colombia.

Roth, André-Noël (2009). La evaluación de políticas públicas en Colombia: una mirada crítica a partir de las prácticas evaluativas oficiales de los programas de la “Red de Apoyo Social”. En Revista Reforma y Democracia, 45. Caracas: Centro Latinoamericano de Administración para el Desarrollo CLAD. [En línea] Recuperado: (6 de septiembre, 2010) Disponible en: <http://www.clad.org/portal/publicaciones-del-clad/revista-clad-reforma-democracia/articulos/045-octubre2-2009/roth>

Sabatier, P. (Ed.) (2007). Theories of the Policy Process (2d ed.). Boulder: University of California, Westview Press.

Salazar, C. (1995). Las políticas públicas. Bogotá: Pontificia Universidad Javeriana. Facultad de Ciencias Jurídicas.

Shaker, P. (1990) “The metaphorical journey of evaluation theory”, International Journal of Qualitative Studies in Education ,Vol. 3, Iss. 4, 1990, 355-363.

Stake, R.E., (1986), “Quieting Reform”, Urbana/Chicago: University of Illinois Press.

Stufflebeam, D., Shinkfield, A. (1981), "Meta-Evaluation: Concepts, Standards and Uses" en R. Berk., ed., Educational Evaluation Methodology: The State of the Art, Johns Hopkins. - (1987),

\_\_\_\_\_ (2001) “The Metaevaluation Imperative”, American Journal of Evaluation, Vol. 22, No. 2, pp. 183–209.

Scriven, M., (2011), Evaluation Thesaurus, Newbury Park: Sage.

\_\_\_\_\_ (1969) "An introduction to meta- evaluation", Educational Product Report 2 (5): 36-38.

\_\_\_\_\_ (2011) "Evaluating evaluations: A meta/evaluation checklist", Claremont Graduate University.

The Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards (2nd ed.). Thousand Oaks, CA: Sage.

United Nations Evaluation Group (2005) "Standards for Evaluation in the UN System".

Toledo, R. (2005). Fortalecimiento de la función evaluación de los países de América del Sur. Informe técnico. Eficiencia y eficacia de la Acción Gubernamental: un Análisis Comparativo de Sistemas de Evaluación. Brasilia: Cooperación técnica Banco Interamericano de Desarrollo e Instituto de Investigación Económica Aplicada.

Villareal, J. (2007, diciembre). Colombia y el Sistema de Monitoreo y Evaluación, SINERGIA. CLAD, Banco Mundial. Documentos Estudios de Casos del CLAD. No. 3.

Weiss, C. (1999). Investigación evaluativa. Métodos para determinar la eficiencia de los programas de acción. Quinta reimpresión. México: Trillas.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). The program evaluation standards: A guide for evaluators and evaluation users (3rd ed.). Thousand Oaks, CA: Sage

Zavakou, Alkistis, (2010?) "An Evaluation of the CCT Program Familias en Acción on Grade Failure and School Drop-out Rates" European University Institute Florence

